

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 10:17:46 ; Search time 69.28 Seconds

(without alignments)
2100.011 Million cell updates/sec

Title: US-09-819-946-2

Sequence: 1 MLCYARLVGQLLISCCWA.....NSTEHQASIDYTRRCGST 841

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

SPTREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteria:*
17: sp._archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3365.5	74.9	842	11 099PG5	099PG5 mus musculus
2	3361.5	74.8	840	11 0920R8	0920R8 rattus norv
3	3355.5	74.7	842	11 0925I5	0925I5 mus musculus
4	3348.5	74.5	842	11 0923J9	0923J9 mus musculus
5	3324.5	74.0	842	11 099PG6	099PG6 mus musculus
6	1370	30.5	843	11 0920R7	0920R7 rattus norv
7	1363	30.3	843	11 0925I4	0925I4 mus musculus
8	1350	30.0	843	11 0923J8	0923J8 mus musculus
9	1170	26.0	858	11 0923K1	0923K1 rattus norv
10	1163	25.9	858	11 0925D8	0925D8 mus musculus
11	1159	25.8	858	11 0925A4	0925A4 mus musculus
12	1147	25.5	858	11 0925D9	0925D9 mus musculus
13	1147	25.5	858	11 091VA4	091VA4 mus musculus
14	1141	25.4	858	11 0923K0	0923K0 mus musculus
15	1060.5	23.6	940	13 073635	073635 fugu rubrip
16	1054	23.5	864	13 073637	073637 fugu rubrip

17	1053	23.4	880	13 073639	073639 fugu rubrip
18	1052.5	23.4	940	13 090WL6	090WL6 sparus aura
19	1021	22.7	848	13 093553	093553 carassius a
20	1014.5	22.6	868	13 073636	073636 fugu rubrip
21	996	22.2	856	13 073638	073638 fugu rubrip
22	975	21.7	877	13 09PW88	09PW88 carassius a
23	949.5	21.1	875	13 073640	073640 fugu rubrip
24	940.5	20.9	844	13 093552	093552 carassius a
25	928.5	20.7	912	11 070410	070410 mus musculus
26	745.5	16.6	855	11 070409	070409 mus musculus
27	737.5	16.4	1156	13 0980C6	0980C6 gallus gall
28	737.5	16.4	1188	13 0980C5	0980C5 gallus gall
29	737.5	16.4	1242	13 0980C4	0980C4 gallus gall
30	732	16.3	1199	13 09EPV6	09EPV6 mus musculus
31	703	15.6	1218	13 090PE3	090PE3 oncorhynch
32	695.5	15.5	983	11 062516	062516 rattus norv
33	692	15.4	779	11 035269	035269 rattus norv
34	687.5	15.3	872	4 09H3N6	09H3N6 homo sapien
35	687	15.3	879	11 09QYS2	09QYS2 mus musculus
36	638	14.2	977	13 09PWE1	09PWE1 ictalurus p
37	615.5	13.7	850	11 035189	035189 mus musculus
38	599	13.3	866	11 035268	035268 rattus norv
39	586.5	13.1	458	13 093555	093555 carassius a
40	586	13.0	803	11 035191	035191 mus musculus
41	561	12.5	852	11 035192	035192 mus musculus
42	547.5	12.2	768	11 035266	035266 rattus norv
43	542	12.1	695	11 035272	035272 rattus norv
44	535	11.9	802	11 035271	035271 rattus norv
45	531.5	11.8	408	13 093558	093558 carassius a

ALIGNMENTS

RESULT	ID	1	PRELIMINARY;	PRT;	842 AA.
099PG5	099PG5	01-JUN-2001 (TREMBLREL. 17, Created)			
AC	099PG5	01-JUN-2001 (TREMBLREL. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLREL. 19, Last annotation update)				
DE	POTATIVE SWEET TASTE RECEPTOR T1R1 (FRAGMENT).				
GN	GPR70.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	11				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=129P3/J;				
RX	MEDLINE=21030739; PubMed=11178737;				
RA	Li X., Inoue M., Reed D.R., Hugue T., Puchalski R.B., Tordoff M.G.,				
RA	Nimolija Y., Beauchamp G.K., Bachmanov A.A.;				
RT	"High-resolution genetic mapping of the saccharin preference locus				
RT	(Sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to				
RT	mouse distal chromosome 4."				
RL	Enbl. Genome 12:13-16(2001).				
DR	EMBL; AF301162; AAK0702.1;				
DR	MGI; MGI:1340021; Gpr70.				
DR	InterPro; IPR001828; ANF_receptor.				
DR	InterPro; IPR000345; Cytc_heme_bind.				
DR	InterPro; IPR000337; GPCR_Mgr.				
DR	Pfam; PF00003; 7tm_3; 1.				
DR	Pfam; PF01094; ANF_receptor; 2.				
DR	PRINTS; PR00248; GPCRMR.				
DR	PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.				
DR	PROSITE; PS50259; G_PROTEIN_RECPEP_F3_4; 1.				
KW	Receptor.				
FT	NON_TER	842	842		
SO	SEQUENCE	842 AA;	93428 MW;	39739A2FF482D33F CMC64;	

Query Match

74.9%; Score 3365.5; DB 11; Length 842;

Best Local Similarity 73.9%; Pred. No. 5e-275;
Matches 622; Conservative 86; Mismatches 136; Indels 1; Gaps 1;

```

OY 1 MLCAR-LVGIQLILSCMAFACHSTESSPDFTLPBGDLVLGAPLPHSGCQVRRPREV 59
DB 1 MLEMAHLILSLQVLAVACMAFSCQRTSSPGFSLPGDFLLGLSLHADLCQVRRPLV 60
OY 60 TLCDSCSFENHGYHLFQAMRLGVEEINNSTALLPNTITGLGYOLYDVCSSANVATRLV 119
DB 61 TSCDSDSFNGNGYHLFQAMRLGVEEINNSTALLPNTITGLGYOLYDVCSSANVATRLV 120
OY 120 SLPGOHIELOGLLHYSPVLAVIGPDSFNRAATALLSPFLVPMISYASSSETLSVK 179
DB 121 AQQGTHLEMOHDLRNHSSKVALIGPDNTDAVTTAALISPLMLPVSEASSVTLSSK 180
OY 180 RQPSFELTIPNDKYQVEHMLLQKFGMTSLVSSSDYQOLGVALENATOGTICI 239
DB 181 RFPSPFLRTIPSDKYQVEHMLLQKFGMTSLVSSSDYQOLGVALENATOGTICI 240
OY 240 AFKDIPEFSAQYDERMOCIMRHLAQAGATVVVVFSSROLARFEFESVLTMLTGKVMVA 299
DB 241 AFKDVVPLSAQAGDPRMORMRLARARTVVVVFSSNRHLAAGVFRSVLANLTGKVMVA 300
OY 300 SEAMALSRHTIGVPGIQRIGVNLGVALIQRAVGLKAFEEAVARADKEAPRCHKSGWCS 359
DB 301 SEDMAISYITVNVPGIIGIGTVLGVALIOQROYGLKEFEESYQAVTGAFCRCPGSGWG 360
OY 360 SNOLECOEAFHMPKIKAFSSANVAYRAYVAHGLHOLLGCGASGRGVVW 419
DB 361 TNOLECOEAFHMPKIKAFSSANVAYRAYVAHGLHOLLGCGASGRGVVW 420
OY 420 QLEOIHKVFLLHKTVAFNDNRDPLSSYNIITLDMNGPKMTFTVLGSGTSPVOLNTN 479
DB 421 QLOOQIYKVFLLHKTVAFNDNRDPLSSYNIITLDMNGPKMTFTVLGSGTSPVOLNTN 480
OY 480 ERTKMGKDNQVPKSVSCSCLLEGHQRYVTGFHHCCEVCVCGAGTFLNKSDDLRCQPC 539
DB 481 KRTKMGKDNQVPKSVSCSCLLEGHQRYVTGFHHCCEVCVCGAGTFLNKSDDLRCQPC 540
OY 540 GKEEWAPEGSQCFRTVYVFLREHNSVTLAANTLILLLGLTAGLFAHMLDPVWVS 599
DB 541 GKEEWAPEGSQCFRTVYVFLREHNSVTLAANTLILLLGLTAGLFAHMLDPVWVS 600
OY 600 AGRLCFLMLGSLAAGSGSLVGFEPTRPACILRQALFALFTIFLSCLTVRSFQIIT 659
DB 601 AGRLCFLMLGSLAAGSGSLVGFEPTRPACILRQALFALFTIFLSCLTVRSFQIIT 660
OY 660 EKESTKVPFTFYHAMVONHAGLFLVMSAQLLICTLWLVMTPLPAREYORPHLMYE 719
DB 661 EKESTKVPFTFYHAMVONHAGLFLVMSAQLLICTLWLVMTPLPAREYORPHLMYE 720
OY 720 CFEETSLIGFLIATVYVFLAVIGPDSFNRAATALLSPFLVPMISYASSSETLSVK 779
DB 721 CFEETSLIGFLIATVYVFLAVIGPDSFNRAATALLSPFLVPMISYASSSETLSVK 780
OY 780 ASVYDGKYLPAANMAGSLSSGEGGYFLPKCYVILCRPDINSTEHFOASIODYTRRCG 839
DB 781 ASVYDGKYLPAANMAGSLSSGEGGYFLPKCYVILCRPDINSTEHFOASIODYTRRCG 840
OY 840 ST 841
DB 841 TT 842

```

RESULT 2

O920R8 PRELIMINARY: PRT: 840 AA.

AC O920R8: 01-MAY-1999 (TREMUREL. 10, Created)
DT 01-MAY-1999 (TREMUREL. 10, Last sequence update)
DE 01-DEC-2001 (TREMUREL. 19, Last annotation update)
OS PUNATIVE TASTE RECEPTOR TRI (FRAGMENT).
Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR;
RA MEDLINE=99159821; PubMed=10052456;
RX HOON M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J.,
RA Zuker C.S.;
RT "putative mammalian taste receptors: a class of taste-specific GPCRs
with distinct topographic selectivity";
RL Cell 96:541-551(1999).
DR EMBL: AF127389; AAD18069.1;
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF000003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 2.
DR PRINTS: PR00248; GPCRMR.
DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 2.
DR Receptor.
FT NON_TER 840
SQ SEQUENCE 840 AA; 93496 MW; 1FCFB7EFC6B45DB CRC64;

Query Match 74.8%; Score 3361.5; DB 11; Length 840;
Best Local Similarity 74.1%; Pred. No. 1.e-274;
Matches 623; Conservative 81; Mismatches 136; Indels 1; Gaps 1;

```

OY 1 MLCAR-LVGIQLILSCMAFACHSTESSPDFTLPBGDLVLGAPLPHSGCQVRRPREV 60
DB 1 MLEMAHLILSLQVLAVACMAFSCQRTSSPGFSLPGDFLLGLSLHADLCQVRRPLV 59
OY 61 TLCDSCSFENHGYHLFQAMRLGVEEINNSTALLPNTITGLGYOLYDVCSSANVATRLV 120
DB 60 SCDRPSDFNGNGYHLFQAMRLGVEEINNSTALLPNTITGLGYOLYDVCSSANVATRLV 119
OY 121 LFGOHIELOGLLHYSPVLAVIGPDSFNRAATALLSPFLVPMISYASSSETLSVK 180
DB 120 LOGPHRIEIOKDLRNHSSKVALIGPDNTDAVTTAALISPLMLPVSEASSVTLSSK 179
OY 181 OYPSFLRTIPNDKYQVEHMLLQKFGMTSLVSSSDYQOLGVALENATOGTICI 240
DB 180 KFPSPFLRTIPNDKYQVEHMLLQKFGMTSLVSSSDYQOLGVALENATOGTICI 239
OY 241 FDIPEFSAQYDERMOCIMRHLAQAGATVVVVFSSROLARFEFESVLTMLTGKVMVA 300
DB 240 FDIPEFSAQYDERMOCIMRHLAQAGATVVVVFSSROLARFEFESVLTMLTGKVMVA 299
OY 301 EAMALSRHTIGVPGIQRIGVNLGVALIQRAVGLKAFEEAVARADKEAPRCHKSGWCS 360
DB 300 EDMAISYITVNVPGIIGIGTVLGVALIOQROYGLKEFEESYQAVTGAFCRCPGSGWG 359
OY 361 NOLCECOEAFHMPKIKAFSSANVAYRAYVAHGLHOLLGCGASGRGVVW 419
DB 360 NOLCECOEAFHMPKIKAFSSANVAYRAYVAHGLHOLLGCGASGRGVVW 420
OY 421 LEOIHKVFLLHKTVAFNDNRDPLSSYNIITLDMNGPKMTFTVLGSGTSPVOLNTN 479
DB 420 LEOIHKVFLLHKTVAFNDNRDPLSSYNIITLDMNGPKMTFTVLGSGTSPVOLNTN 479
OY 481 TKIOMHGKDNQVPKSVSCSCLLEGHQRYVTGFHHCCEVCVCGAGTFLNKSDDLRCQPC 540
DB 480 TKIOMHGKDNQVPKSVSCSCLLEGHQRYVTGFHHCCEVCVCGAGTFLNKSDDLRCQPC 539
OY 541 KEWAPEGSQCFRTVYVFLREHNSVTLAANTLILLLGLTAGLFAHMLDPVWVS 600
DB 540 KEWAPEGSQCFRTVYVFLREHNSVTLAANTLILLLGLTAGLFAHMLDPVWVS 599
OY 601 GGRLCFLMLGSLAAGSGSLVGFEPTRPACILRQALFALFTIFLSCLTVRSFQIIT 660
DB 600 GGRLCFLMLGSLAAGSGSLVGFEPTRPACILRQALFALFTIFLSCLTVRSFQIIT 659
OY 661 KFTKVPFTFYHAMVONHAGLFLVMSAQLLICTLWLVMTPLPAREYORPHLMYE 720

```



```

Db 660 KESTKVPFTFYHTAQNNGAGLFVIVSSTVHLICTLMTVMWTPRPREYQGFPHLVILEC 719
Qy 721 TENNSLGFILAFLYNGLISTISAFACSYLGKDLPENYNKACVTSLEFNFSWIAFETM 780
Db 720 TENNSVGFELVAFHNLISTFVCSYLGKELPENYNKACVTSLEFNFSWIAFETM 779
Qy 781 SVYDGKYLPAANMAGLSLSSGFGYFLPKCYIILCRPDNSTEHFQASIODYTRRCGS 840
Db 780 SITVGSFLPANNVLAGLITLSGFSGYFLPKCYIILCRPELNTNHEHQASIODYTRRCGT 839
Qy 841 T 841
Db 840 T 840

RESULT 3
Q92515 PRELIMINARY; PRT; 842 AA.
AC Q92515;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CANDIDATE TASTE RECEPTOR T1RL.
GN T1RL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21219400; PubMed=11319557;
RA Montmayeur J.P., Liberles S.D., Matsuami H., Buck L.B.;
RT "A candidate taste receptor gene near a sweet taste locus.";
RL Nat. Neurosci. 4:492-498(2001).
EMBL: AF337040; AAK39437.1; -.
KM Receptor.
SQ SEQUENCE 842 AA; 93471 MW; FDBAOCCLBDA5DEE7 CRC64;

Query Match 74.7%; Score 3355.5; DB 11; Length 842;
Best Local Similarity 73.6%; Pred. No. 3.5e-274;
Matches 620; Conservative 87; Mismatches 134; Indels 1; Gaps 1;

Qy 1 MLCSTAR-LVGLQLLISCCNAFACHSTESSPDTLPDGYLLAGLFPPLHSGCLOYRHPREV 59
Db 1 MLEFAAHLLSLQLAVAYCMAFSCORTSSPGFSLPGDFLLAGLFSHADCLQVRHRLV 60
Qy 60 TLCDRSCSFNKGHYHLFOAMRLGVEEINNSFALLPNTITLGOLYDVCSDSANVATLRLV 119
Db 61 TSCDRSDFNKGHYHLFOAMRLGVEEINNSFALLPNTITLGELYDVCSSSNVATLRLV 120
Qy 120 SLPGOHIELOGDILHSPVLAIVGPDSTNRATTAALSPFLVPMISYASSETLSVK 179
Db 121 AAGCTGHEMQRDLRNHSSKVALIGPDNDHATTAALSPFLPLVSYASSVILSGK 180
Qy 180 ROYPSFLRTIPNDKYOVETWVLLLOKFGWTWISLVGSSDDYGOLGVQALENQATGOGICI 239
Db 181 RKPSFLRTIPDSKYOVETWVLLLOKFGWTWISLVGSGDYGOLGVQALENLATPRGICV 240
Qy 240 AFKQIMPFSAOVDEROQCLMRHLAOGATVVVVFSSROLARFEEFVLTNLGKMWVA 299
Db 241 AFDVVLVLSAOGDPKQMRMLRLARARTVYVVFSSNRHLAGVFFRSVVLNLGKMWIA 300
Qy 300 SEANALSRHITGVGIGRIGAVLGAIOKRAVPLGAKFEENARADREAPRCHGKSWCS 359
Db 301 SEDMAISTYITNVPGIGIGIVLGAIOQROVPGLKEFEESYVOAVMGAPRTCEGSMCG 360
Qy 360 SNQCRCCQAFMATMTRKIAFSMSAYNAYRAYVAHGHOLGASGCSGRYVPM 419
Db 361 TNOICRCHAFETWVNMDELGAFSMSAAYNAYEAVVAHGHOLLGCTGCTGARGPYVPM 420

```

```

Qy 420 QLLQEIHKVHFLHKDYAFANDNRDPLSSSYNIAMDNMGPKWTFVLGSSWSPVOLNIN 479
Db 421 QLLQOIYKVNFLHKHTYAFNDKDPGLGYDILAMDNMGPKWTFVLGSSWSPVOLNIN 480
Qy 480 ETKIOWHGKDNQVPSKSSDCLBGNRVYTGPHHCCEYPCGAGTFLNKSLEYRCOPC 539
Db 481 KTKIOWHGKDNQVPSKSSDCLBGNRVYTGPHHCCEYPCGAGTFLNKSLEYRCOPC 540
Qy 540 GKEENAPGSGOTCEPRTVFVLALREHNTSWVLANTLILLLILGAGLFAMHLDPPVRS 599
Db 541 GTEENAPGSSACFSRTVEFGWHEPISVLDAANTLILLLILGAGLFAMHLDPPVRS 600
Qy 600 AGGRICFLMLGSLAAGSGSLYGFGEPTRPACLLRQALFALGFTFLSCLTVRSQOLIYI 659
Db 601 AGGRICFLMLGSLAAGSGSLYGFGEPTRPACLLRQALFALGFTFLSCLTVRSQOLIYI 660
Qy 660 KESTKVPFTFYHTAQNNGAGLFVIVSSTVHLICTLMTVMWTPRPREYQGFPHLVILEC 719
Db 661 KESTKVPFTFYHTAQNNGAGLFVIVSSTVHLICTLMTVMWTPRPREYQGFPHLVILEC 720
Qy 720 TENNSLGFILAFLYNGLISTISAFACSYLGKDLPENYNKACVTSLEFNFSWIAFETM 779
Db 721 TENNSVGFELVAFHNLISTFVCSYLGKELPENYNKACVTSLEFNFSWIAFETM 780
Qy 780 ASYVDGKYLPAANMAGLSLSSGFGYFLPKCYIILCRPDNSTEHFQASIODYTRRCG 839
Db 781 SITVGSFLPANNVLAGLITLSGFSGYFLPKCYIILCRPELNTNHEHQASIODYTRRCG 840
Qy 840 ST 841
Db 841 TT 842

RESULT 4
Q92309 PRELIMINARY; PRT; 842 AA.
AC Q92309;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE TASTE RECEPTOR T1RL.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
RA Zuker C.S.;
RT "Mammalian Sweet Taste Receptors.";
RL Cell 0:0-0(2001).
EMBL: AY032622; AAK51603.1; -.
KM Receptor.
SQ SEQUENCE 842 AA; 93455 MW; D4D2CE90959E991A CRC64;

Query Match 74.5%; Score 3348.5; DB 11; Length 842;
Best Local Similarity 73.5%; Pred. No. 1.3e-273;
Matches 619; Conservative 87; Mismatches 135; Indels 1; Gaps 1;

Qy 1 MLCSTAR-LVGLQLLISCCNAFACHSTESSPDTLPDGYLLAGLFPPLHSGCLOYRHPREV 59
Db 1 MLEFAAHLLSLQLAVAYCMAFSCORTSSPGFSLPGDFLLAGLFSHADCLQVRHRLV 60
Qy 60 TLCDRSCSFNKGHYHLFOAMRLGVEEINNSFALLPNTITLGOLYDVCSDSANVATLRLV 119
Db 61 TSCDRSDFNKGHYHLFOAMRLGVEEINNSFALLPNTITLGELYDVCSSSNVATLRLV 120
Qy 120 SLPGOHIELOGDILHSPVLAIVGPDSTNRATTAALSPFLVPMISYASSETLSVK 179
Db 121 AAGCTGHEMQRDLRNHSSKVALIGPDNDHATTAALSPFLPLVSYASSVILSGK 180
Qy 180 ROYPSFLRTIPNDKYOVETWVLLLOKFGWTWISLVGSSDDYGOLGVQALENQATGOGICI 239

```

```

Db 181 RRFPSFLRTIPSDKYQVEIYVRLLOSFGVWVMSLVGSYDYDQGLGVALLEELATRGICV 240
QY 240 AKKDIMPESAOYGDERMOCIMRLHQAQATVYVSSRQLAVFPESVLTMLTGKWWA 299
Db 241 AKKDVPISAOQGDPRMOMMLRLARFTVVVFNRLAVVFNRYVLAATLMLTGKWWA 300
QY 300 SEAMALSRHITGVPGIORIGVAVIAOKRAVPGKAFEEAARADKEAPRCHGSMCS 359
Db 301 SEDMAISITVTNVPGIIGITVLGVALIOORQVPGLKEFEESVYQAVMGAPRCPESSMWC 360
QY 360 SNOLORECOAFNAHMPKIKAFSSMSATNAYAVAVAHGLHQLGCSAGSCRGVRYPM 419
Db 361 TNOLORECAFTTMMPELGAFSMSAAYVAVAVAHGLHQLGCTSGTCARGPVYPM 420
QY 420 OLLEOIHVHFLHDKDTVAFNDNRDPLSSYNTIAMDNGPKFTFYLGSSVSPVOLNTN 479
Db 421 QLLQOITKYVNFLLHKKTYAFDDKDPGLGYDIIAMDNGPEMTFEYIGSASLSPVHLDTN 480
QY 480 ETKIQMHGKDNQVPRKSVCSDDLGGHQRVVTGPHHCCFECVPCGAGTFPLNKSDLYRCOPC 539
Db 481 KTKIQMHGKNNQVPRSVCTRDCLGGHRLVMGSHHCCFECMPCGAGTFPLNTSELHTCOPC 540
QY 540 GKEENAPBESSQCFPRTVYFLALREHTSVWVLAANTLLLLLTGTAFLFAMHLDPVYVS 599
Db 541 GKEENAPBESSACFSRTVYFLFWMHBPISLVLAANTLLLLLTGTAFLFAMHLDPVYVS 600
QY 600 AGGRICFLMLGSLAAGSGSLYGFEGEPTRPACILROALFALFTFPLSCLYVRSQOLIT 659
Db 601 AGGRICFLMLGSLVAGSGSLYFEGKPTVPACILROALFALFTFPLSCLYVRSQOLIT 660
QY 660 FKFSTRKVPFTYHAWVONHAGLFWMISSAOOLICLTWLVVTPPLPAREYORFPHVME 719
Db 661 FKFSTRKVPFTYHAWVONHAGLFWMISSAOOLICLTWLVVTPPLPAREYORFPHVME 720
QY 720 CRETNSIGFTILAFVNGLSISAFSCSYGKDLPEVYNNAKCVTSLSLFENFSWIAFTT 779
Db 721 CRETNSVGLVFAFNHILISITFVCSYIGKELPENYNNAKCVTSLSLFENFSWIAFTT 780
QY 780 ASVYDGKYLPAANMAGLSLSSGFGGYFLPKCYVILCRPDNSTEHFOASIODYTRRCG 839
Db 781 SSITGGSYLPANVNLAGLATLSGFGSGYFLPKCYVILCRPELNTEHFOASIODYTRRCG 840
QY 840 ST 841
Db 841 TT 842

```

```

DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm.3.1.
DR PRINTS: PR00248; GPCRMR.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
DR PROSITE: PS00259; G_PROTEIN_RECEP_F3_4; 1.
KW Receptor.
FT NON_TER
SEQUENCE 842 AA; 93425 MW; 58826C43F5D0352E CRC64;

Query Match 74.0%; Score 3324.5; DB 11; Length 842;
Best Local Similarity 73.0%; Pred. No. 1.4e-271;
Matches 615; Conservative 89; Mismatches 137; Indels 1; Gaps 1;

QY 1 MLCSTAR-LVGLQOLISCOMAFACSTESSPDFTLPQDVLGAGLPRLSSGLOVHRREV 59
Db 1 MTFMAHLLSLQLAVAYCMASFSCORTSSPEFSLPGBDLGSLHADCLQVHRRLV 60
QY 60 TLCDRSCSFNEGHYLFQAMRLGVEEINNSTALLPNTILGYOLYDVCSDSANVATRLV 119
Db 61 TSCDRSDSPNGHYHLPQAMRFTVEEINNSTALLPNTILGYELDYDVCSSNVATRLV 120
QY 120 SLPGCHHIFLQDGLHYSTVYAVIGPDSNRAATTAALLSPFLVPMISYAASETLSVK 179
Db 121 AQQGCHLEMOQDLRNHSSKVVYALIGPMDTHAVYTAALLSPFLVPMISYAASETLSVK 180
QY 180 RQYPSFLRTIPNDKYQVEIYVRLLOSFGVWVMSLVGSYDYDQGLGVALLEELATRGICV 239
Db 181 RKFPSFLRTIPSDKYQVEIYVRLLOSFGVWVMSLVGSYDYDQGLGVALLEELATRGICV 240
QY 240 AKKDIMPESAOYGDERMOCIMRLHQAQATVYVSSRQLAVFPESVLTMLTGKWWA 299
Db 241 AKKNVYPLSAQGDPRMOMMLRLARFTVVVFNRLHLDGFFRSVYLAATLMLTGKWWA 300
QY 300 SEAMALSRHITGVPGIORIGVAVIAOKRAVPGKAFEEAARADKEAPRCHGSMCS 359
Db 301 SEDMAISITVTNVPGIIGITVLGVALIOORQVPGLKEFEESVYQAVMGAPRCPESSMWC 360
QY 360 SNOLORECOAFNAHMPKIKAFSSMSAAYVAVAVAHGLHQLGCSAGSCRGVRYPM 419
Db 361 TNOLORECAFTTMMPELGAFSMSAAYVAVAVAHGLHQLGCTSGTCARGPVYPM 420
QY 420 OLLEOIHVHFLHDKDTVAFNDNRDPLSSYNTIAMDNGPKFTFYLGSSVSPVOLNTN 479
Db 421 QLLQOITKYVNFLLHKKTYAFDDKDPGLGYDIIAMDNGPEMTFEYIGSASLSPVHLDTN 480
QY 480 ETKIQMHGKDNQVPRKSVCSDDLGGHQRVVTGPHHCCFECVPCGAGTFPLNKSDLYRCOPC 539
Db 481 KTKIQMHGKNNQVPRSVCTRDCLGGHRLVMGSHHCCFECMPCGAGTFPLNTSELHTCOPC 540
QY 540 GKEENAPBESSQCFPRTVYFLALREHTSVWVLAANTLLLLLTGTAFLFAMHLDPVYVS 599
Db 541 GKEENAPBESSACFSRTVYFLFWMHBPISLVLAANTLLLLLTGTAFLFAMHLDPVYVS 600
QY 600 AGGRICFLMLGSLAAGSGSLYGFEGEPTRPACILROALFALFTFPLSCLYVRSQOLIT 659
Db 601 AGGRICFLMLGSLVAGSGSLYFEGKPTVPACILROALFALFTFPLSCLYVRSQOLIT 660
QY 660 FKFSTRKVPFTYHAWVONHAGLFWMISSAOOLICLTWLVVTPPLPAREYORFPHVME 719
Db 661 FKFSTRKVPFTYHAWVONHAGLFWMISSAOOLICLTWLVVTPPLPAREYORFPHVME 720
QY 720 CRETNSIGFTILAFVNGLSISAFSCSYGKDLPEVYNNAKCVTSLSLFENFSWIAFTT 779
Db 721 CRETNSVGLVFAFNHILISITFVCSYIGKELPENYNNAKCVTSLSLFENFSWIAFTT 780
QY 780 ASVYDGKYLPAANMAGLSLSSGFGGYFLPKCYVILCRPDNSTEHFOASIODYTRRCG 839
Db 781 SSITGGSYLPANVNLAGLATLSGFGSGYFLPKCYVILCRPELNTEHFOASIODYTRRCG 840
QY 840 ST 841

```

Db 841 TT 842

RESULT 6

0920R7 PRELIMINARY; PRT; 843 AA.

AC 0920R7; 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PUTATIVE TASTE RECEPTOR TR2 (FRAGMENT).

OS Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

RM [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MISTAR;

RX MEDLINE=99159821; PubMed=10052456;

RA Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J., Zuker C.S.,

RT "Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity."

RL Cell 96:541-551(1999).

DR EMBL; AF127390; AAD18070.1; -.

DR InterPro: IPR001828; ANF_receptor.

DR InterPro: IPR000337; GPCR_Mgr.

DR Pfam; PF00003; 7tm_3; 1.

DR Pfam; PF01094; ANF_receptor; 2.

DR PRINTS; PR00248; GPCRMR.

DR PROSITE; PS50259; G_PROTEIN_RECPEP_F3_4; 2.

KW Receptor.

FT NON_TER

SO SEQUENCE 843 AA; 95799 MW; D23AC22D1E04988 CRC64;

Query Match 30.5%; Score 1370; DB 11; Length 843;
Best Local Similarity 37.8%; Pred. No. 1e-106;
Matches 318; Conservative 145; Mismatches 316; Indels 62; Gaps 22;

QY 31 DFTLPDYLGLGFLPHSGCLOYRHRP--EYTLCDRSCSFNEGYHLFOAMRLGYEINN 88
DB 29 DFTLAGDYLLGLGFLPHANVKSISLSYLQVPCN-EFTMKVLYGLNMQAMFAVEINN 87
QY 89 SFALLPNTTIGLYOLVDCSDSANVATRLVLSLPGNHLELOGDL-----HSPYPL 141
DB 88 CSSLPRLGVLLGEMVDVCLSNNH-----PGLYFLADDDDLRLPLIKDYQYPTPHV 139
QY 142 AVIGPDSNRATTAALLSPFLVPMISYASSETLSVKRQPSFLRTTIPNDKYQVETVL 201
DB 140 AVIGPDSNRSATTVSNILSHFLIPQITYSALDKLRKHNFPMSLRTVPSATHHIEAMV 199
QY 202 LLOKRGWTVISLVGSSDDYQGLGVQALENQAT--GGGICIAFQDIMP--SAQV--GDERM 256
DB 200 LAMVHQQMMWIVLVDDDYGRNSHLLSQRLTKTSDICIAFOEVLPIRPSSQVMRSEOR 259
QY 257 QC--LMRHLAOGATVYVVSRLARVFEESVLTNLGKTVWASEALRHTITGVG 314
DB 260 QLDNLTDLKRTSARVYVVSPELSLYFHEVLRNFGFVWASEALIDPVLHNTLE 319
QY 315 IORIGMVLGAIOAKRVAGLKAFAEAYARADKEA--PRPCHKSWCSSNOLCRECOAFMAH 373
DB 320 LRHTGTPLGVTIQRVISIFQSFQF--VRDRKGRVPVPTNLTTCNDDCAC----- 370
QY 374 TMRPKAKS-----MSS-----AYNAYRAVYAVAHGLHQLGACAGACSGRVYWMOLLEQTH 426
DB 371 -LNTTKSFNNILISGERVYVSVAVAHARHLRGLGCMNVKCTKQVYPMOLIRETW 429
QY 427 KHFPLHKQTVAFNDNRDPLSSYNTIAMDNGPKMTFTVLGSGTSPVQLNINE--TKIQM 485
DB 430 HNFPLLLGRLEFFDOGGMPMLLDILIQOMDLSQNPFOSIAS--YSPISKRLTYINNVSW 487
QY 486 HGKDQVPAKSVCSDDLSEGHQVTVTGHNHCEPCVPCGAGTFLNKS-DLYRCOPCGKEEM 544

Db 488 YTPNNTVPVSMCSKSCOPGOMKRSVGLHPCCFELCDMPGTLYNRSADFNCLSCPSMW 547
QY 545 APESQTCFPRRTVYFIALRHTSMV--LLAA-----NTLLLLLLGLAGLFAMHLDPVVR 538
DB 548 SYKNDITTCFORPRTFLEMHVPPRIIVAILAALFAEFSTLALF-----IFMRHQPDMVR 601
QY 599 SAGRCFLMGLSLAAGSGSLYGFEGEPTRPACLLQVAFALGFTIFLSCGLVRSPLII 658
DB 602 SAGCPMCFMLVLLAFMGVAPVYVGPVYVSCFCQAFVTFVFSICLSCTITRSDIYC 661
QY 659 IFKSTKVPTEFYHAWYONHAGLFVMISSAQLLICLTWLVVTPLP--AREYORFPLHY 717
DB 662 VFKMARLRPSAYEFMRHGRPYFAFTRIKALVAGNMALATTIPDICTDDDDNIMI 721
QY 718 LECTETNSLGFILAFLYNLGLSLAFACSTYLGKDLDEBNENAKCVFSLFNVSMAF 777
DB 722 LSCHPVYRNGLLFNTSMDDLISVLGFSFAYMGKELPTNYNEAKFITLSMTFSSTISLC 781
QY 778 TTSAYVDGKYLPAANMAG-TSLSSGFGSELPKCVILCRPDLNSTERFOASIDQYTR 836
DB 782 TFKSVHDGVLVITMDLVTVLNFALGL-GYFSPKCMILFPERNTSAVFSMIGQYTM 840
QY 837 R 837
DB 841 R 841

RESULT 7

092514 PRELIMINARY; PRT; 843 AA.

AC 092514; 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE CANDIDATE TASTE RECEPTOR TIR2.

GN TIR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RM [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J;

RX MEDLINE=21219400; PubMed=11319557;

RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B., "A candidate taste receptor gene near a sweet taste locus."

RL Nat. Neurosci. 4:492-498(2001).

DR EMBL; AF337041; AAK39438.1; -.

KW Receptor.

SO SEQUENCE 843 AA; 95735 MW; 0543834EA4E7AC2E CRC64;

Query Match 30.3%; Score 1363; DB 11; Length 843;
Best Local Similarity 37.8%; Pred. No. 4e-106;
Matches 316; Conservative 142; Mismatches 325; Indels 54; Gaps 19;

QY 31 DFTLPDYLGLGFLPHSGCLOYRHRP--EYTLCDRSCSFNEH-----GYHLFOAMRLGY 83
DB 29 DFTLAGDYLLGLGFLPHANVKSISLSYLQVPCN-----NEXNMVLYGLNMQAMFAV 82
QY 84 EELNNSTALLPNTTIGLYOLVDCSDSANVATRLVLS-----LPQNHLELOGDLHNSP 138
DB 83 EELNNSSLLPVLGLGEMVDVCLSNNHQPLGYFLSQDDFLP-----ILKDYQYRP 136
QY 139 TVLAVIGPDSNRATTAALLSPFLVPMISYASSETLSVKRQPSFLRTTIPNDKYQVET 198
DB 137 QVAVAYGPDNSRSATTVSNILSYFLVQYTYSAITDKLRKRRFPMLRTPVPSATHHIEA 196
QY 199 MYLLLOKFGWTVISLVGSSDDYQGLGVQALENQATGQ--ICIAFKQDIMP--SAQVED 253
DB 197 MYQLAMVHQQMMWIVLVVDDDYGRNSHLLSQRLTKTGTGICIAFOEVLPIRPENQAVRPE 256

SEQUENCE 858 AA; 94845 MW; 98890DAE75973B80 CRC64;

Query Match

26.0%; Score 1170; DB 11; Length 858;

Best Local Similarity 33.3%; Pred. No. 7.8e-90; Matches 277; Conservative 142; Mismatches 358; Indels 54; Gaps 17;

QY SPDTLPEDYLLAGLPPL--HSGCLOVRHREVTLCDRSCSFNENGYHNLFOAMRLGYEE 85
 DB SQEFAQGDYILGFLPGSTEATLNORPNSILCTR--FSPGLFLMAMKMAVEE 82
 QY INNSTALLPNTTIGYQYDVCSDS-ANYATRLVSLPGQHHIELQDGLHYSPVLAVI 144
 DB INNGSALLPGLRLGDLDTCESEPVYTKKPSLMFAKVGSGSIAAYCNYQYQRYLAVI 142
 QY 145 GPDSTNRAATTAALSPFLVPMISYAASETLSVKROYPSFLRTIPNDKYOVETWLLQ 204
 DB 143 GPHSESLALITGKFPSEFLMPQVSYASMDRLSDRETFPSFRVYPSDRVOLQAVVLLQ 202
 QY 205 KFGWTWISLVGSSDDYGGOLGVALENQATGGGICIAFDIMPFSAQVDE--RMQCLMRH 262
 DB 203 NFSNMWVAALGSDDYREGISIFSGLANSGICIAHGLVPOHDTSGOQLKAVDVLRQ 262
 QY 263 LAQAGATVYVYFSSRQLARVFESVVLNTLTKVWVASEAMALSRHTGVGIORIGVYL 322
 DB 263 VNQSKVQVYVLFASARAVYSLFYSYLHDLSPKYVWASESWLSDLVMTLPRIARVTVL 322
 QY 323 GVALOKRAVPGIKAFEEAYARADKEAPRPCHKSGWSS-----NOLCREQCAF 370
 DB 323 GFLRGALLPEFSHYVETHLAL--AADP---FFCASLMELEDERVMGRCPCQDI 375
 QY 371 MAHT-----MPKLKAFSM-SSAYNAVAVYVAHGLHOLLGASGASGR-YYPMOLLE 423
 DB 376 MLONLSSGLONLSAGOLHNOIFATYAAYVYAQALHNTLOCNVSHCHTSPVQWOLLE 435
 QY 424 OIHKVFLLHKDLYAFENDNRDPLSSYNI IAWDMNGPKWTFVLGSSSTWSPVOLINETKI 483
 DB 436 NMYNMSFRADLTLOFPAKGSVDMEYDLKMMWQSPFVLHTVGTNGT--LOLOHSM 492
 QY 484 OMHGKDNQVPSVSSDCLGHOVRVYGFHNCSEFCVPCGAGTFLNKSNDLYRCOPCGEE 543
 DB 493 YWPG--NOVPVQSCROCKDQVRRVKGFSHCYDVCCKASYRKHHDDFTPCQNDQ 550
 QY 544 WAPESOTCFPRVYVFLALREHTSMVLLAANTLLLLLTAGTFAHMLDTPVVSAGR 603
 DB 551 WSPKSTTCLPRRKFLAMGEPVLSLLLLCLVGLTLALGFLVHNDSPVQASGS 610
 QY 604 L-CF--LMLGSLAAGSGSLYGFGEPTRPACLLRQALFALGFTIFLSCLVRSFOLITF 660
 DB 611 LFCGLICLIGFLC---SVLLFPGRPSSASCLAQOPMAHLPLTGCLSTFLQAAEIFV-- 665
 QY 661 KFSKVPFTFYHAWQNHGAG---LFYMISAAQLLICLTMLVWVTPLPAREYQRFPHLY 716
 DB 666 --ESELPSMANWMLCSYLRGMAWLVVLLATLVEALCALMAFPRPVYVDMOVLPREV 723
 QY 717 MLECTETNSLGFILAFYNGLLISAFACSYLGRDLPENYNAKCVTSLEFNFSWIAF 776
 DB 724 LEHRMNSWVSLGLVHTNNAFLCLGTFPLVOSQGRYRNARGLTPRAMLAYFTIWSF 783
 QY 777 FTTSAYVDGKYLPANNAAGLSLSSGEGYELPKCYVILCRPDINSTENF 827
 DB 784 VPLIANOVAVOPAVOMGALIFCALGILATFHLPKCYVLLMLPELNOEFF 834

RESULT 10

ID Q925D8 PRELIMINARY; PRT; 858 AA.
 AC Q925D8;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PUTATIVE SWEET TASTE RECEPTOR FAMILY 1 MEMBER 3.
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVEV;
 RX MEDLINE=21225287; PubMed=11326277;
 RA Max M., Shanker Y.G., Huang L., Rong M., Campagne F., Liu Z.,
 RA Weinstein H., Damak S., Margolskee R.F.,
 RT "Tas1r3, encoding a new candidate taste receptor, is allelic to the
 RT sweet responsiveness locus Sac."
 RL Nat. Genet. 28:58-63(2001).
 DR EMBL; AF368025; AAK5537.1; .
 KW Receptor.
 SO SEQUENCE 858 AA; 94533 MW; 2012102B55D7E6B0 CRC64;

Query Match 25.9%; Score 1163; DB 11; Length 858;
 Best Local Similarity 33.0%; Pred. No. 3e-89;
 Matches 274; Conservative 150; Mismatches 353; Indels 54; Gaps 16;

QY SPDTLPEDYLLAGLPPLHS--GCLQVRHREVTLCDRSCSFNENGYHNLFOAMRLGYEE 85
 DB SQEFAQGDYILGFLPGSTEATLNORPNSILCTR--FSPGLFLMAMKMAVEE 82
 QY INNSTALLPNTTIGYQYDVCSDS-ANYATRLVSLPGQHHIELQDGLHYSPVLAVI 144
 DB INNGSALLPGLRLGDLDTCESEPVYTKKPSLMFAKVGSGSIAAYCNYQYQRYLAVI 142
 QY 145 GPDSTNRAATTAALSPFLVPMISYAASETLSVKROYPSFLRTIPNDKYOVETWLLQ 204
 DB 143 GPHSESLALITGKFPSEFLMPQVSYASMDRLSDRETFPSFRVYPSDRVOLQAVVLLQ 202
 QY 205 KFGWTWISLVGSSDDYGGOLGVALENQATGGGICIAFDIMPFSAQVDE--RMQCLMRH 262
 DB 203 NFSNMWVAALGSDDYGREGISIFSSLANARICIAHGLVPOHDTSGOQLKAVDVLRQ 262
 QY 263 LAQAGATVYVYFSSRQLARVFESVVLNTLTKVWVASEAMALSRHTGVGIORIGVYL 322
 DB 263 VNQSKVQVYVLFASARAVYSLFYSYLHDLSPKYVWASESWLSDLVMTLPRIARVTVL 322
 QY 323 GVALOKRAVPGIKAFEEAYARADKEAPRPCHKSGWSS-----NOLCREQCAF 370
 DB 323 GFLRGALLPEFSHYVETHLAL--AADP---ARCASLMELEDERVMGRCPCQDI 375
 QY 371 MAHTPKLKAFMSMA-----YNAVAVYVAHGLHOLLGASGAS-REGVYPMOLLE 423
 DB 376 MLONLSSGLONLSAGOLHNOIFATYAAYVYAQALHNTLOCNVSHCHVSEHVLWOLLE 435
 QY 424 OIHKVFLLHKDLYAFENDNRDPLSSYNI IAWDMNGPKWTFVLGSSSTWSPVOLINETKI 483
 DB 436 NMYNMSFRADLTLOFPAKGSVDMEYDLKMMWQSPFVLHTVGTNGT--LOLOHSM 492
 QY 484 OMHGKDNQVPSVSSDCLGHOVRVYGFHNCSEFCVPCGAGTFLNKSNDLYRCOPCGEE 543
 DB 493 YWPG--NOVPVQSCROCKDQVRRVKGFSHCYDVCCKASYRKHHDDFTPCQNDQ 550
 QY 544 WAPESOTCFPRVYVFLALREHTSMVLLAANTLLLLLTAGTFAHMLDTPVVSAG- 602
 DB 551 WSPKSTTCLPRRKFLAMGEPVLSLLLLCLVGLTLALGFLVHNDSPVQASGS 610
 QY 603 RLCP--LMLGSLAAGSGSLYGFGEPTRPACLLRQALFALGFTIFLSCLVRSFOLITF 660
 DB 611 QPCFLICLIGFLC---SVLLFPGRPSSASCLAQOPMAHLPLTGCLSTFLQAAEIFV-- 665
 QY 661 KFSKVPFTFYHAWQNHGAGF---VMISAAQLLICLTMLVWVTPLPAREYQRFPHLY 716
 DB 666 --ESELPSMANWMLCSYLRGMAWLVVLLATLVEALCALMAFPRPVYVDMOVLPREV 723
 QY 717 MLECTETNSLGFILAFYNGLLISAFACSYLGRDLPENYNAKCVTSLEFNFSWIAF 776
 DB 724 LEHRMNSWVSLGLVHTNNAFLCLGTFPLVOSQGRYRNARGLTPRAMLAYFTIWSF 783

```

OY 777 FTASVYDGKYLPPANMMAGLSSSGFGCYLPKCYVILCRPDLNSTEHF 827
DB 784 VPLLANVOAYOPAVOMGAILVICALGILVTFLPKCYVILMLPKLNTQEFF 834

RESULT 11
OY 925M4 PRELIMINARY: PRT: 858 AA.
AC 0925M4:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE TASTE RECEPTOR.
GN TIR3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Sainz E., Koriey J.N., Batley J.F., Sullivan S.L.;
RT "Identification of a novel member of the T1R family of putative taste
RT receptors."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY026318; AAK01937.1;
KW Receptor.
SQ SEQUENCE 858 AA; 94567 MW; A7879C015C0E8474 CRC64;

Query Match 25.8%; Score 1159; DB 11; Length 858;
Best Local Similarity 32.9%; Pred. No. 6,6e-89;
Matches 273; Conservative 150; Mismatches 354; Indels 54; Gaps 16;

OY 29 SPDFLLPGDYLLAGLPLHS---GCLQVRHREPVYLCDSKSCFNEGHYLPQAMRLGVEE 85
DB 26 SOQFAOGDYIIGLPLPGSTEBATLNQAPNSTLCNR--FSPGLFLAMAKMAVEE 82
OY 86 INNSTALLPNTLYGYOLYVCSDS-ANYATRLVLSLPGOHHELOGDILHSPYLAIV 144
DB 83 INNGSALLPGLRLGYDLFTCESEPVYTKSSLMFLAKVSOSIAVCNTQOPRLAVI 142
OY 145 GPDSTNRATTAALSPFLVPMISYAASSETLSVKRQYPSFLRTIPNDKYQVETWVLLQ 204
DB 143 GPHSSELALITGKFFSFLMPQVYSASMDRLSDRETFPSFRTVPSDRVQLOAVYTLQ 202
OY 205 KFGWTWISLVGSDDYGGQLGVALENOATGOCIAFKDIMPESAQVGE--RMQCLMRH 262
DB 203 NFSMNVVAALGSDDDYGRGGLSTFSSLANARGICTAHGELVQHDTSGQOLGKVDLVLRQ 262
OY 263 LAQAGATVVVVFSSRQLARVFESVVLNLTGKVVVASEAMALSRHITGVPGIORIGVL 322
DB 263 VNOSKVVVVFLASARAVYSLEFSYSIHGSLSPKVVVASESWLTSVLVMTLPINARVGYL 322
OY 323 GVAIQRAVPGIKAEFEAYARADKEAPRCHGSCSS-----NOLCRECOAF 370
DB 323 GLQGRALLPERSHYETHAL--AADP---AFCASLMAELDEEHVWGRCPCDDI 375
OY 371 MAHTMPKILKAFMSA-----YNAYRAVYVAHGLHQLGASGACS-GRGYVPQOLLE 423
DB 376 MIONSSGGLQWLSAGOLHJOIFATYAAVSYAVALHNTLQCNVSHCHVSEHYLPQOLLE 435
OY 424 QIHKVFLLHKDTVAFNDRDPLSSYNIITAMDNGPKMTFTVLGSSSTWSPVQNLINETI 483
DB 426 NMYNNSFARDLTLQDAEGNVDMEDLKMVWQSPPTVPLHTVGFNGT---LQLOQSKM 492
OY 484 OMHGNDNOVPSVCSDDLGEHQRYVTGPHHCCFECVPCGAGTFPLKSLYKQCPGKKE 543
DB 483 YWPG--NOVPVQSCROCKDQGVRRYKGFHSCCYDCVCKASYSYRKHDPDFTCTPCNOQO 550
OY 544 WAPESGQCEPFTVYVLLREHTSWVLAANTLILLILGTAGLFWMLHDTFVVSAGS- 602
DB 551 WSPKSTACLPKRPKFLMAGEPVVLSLLLLCTLVGLALALGLSVHMDSPVQVQASGGS 610

```

```

OY 603 RLCE--LMLGSLAAGSGLYGFGEPTRPACLLRQALFALGFTIFLSCLTVRSFQOLIEF 660
DB 611 OCFEGFLICGLFCL---SVLLFPGRRSSASCLAQCPMAHLPLTGCLSTFLQAAEFV-- 665
OY 661 KESTKVPFTFYHAMVONHGGGLF-----VMISAAQOLLCITLVVMTPLPAREQRPFLV 716
DB 666 --ESELPLSMAMWLCGYLGLGAMVLLATVFEALCAMYLTAFPEVVTQWSVLPTEV 723
OY 717 MECEFTNSLGFIAFLVYNGLISAFACSYLGKDLPEVNEAKCYTSLFLEFNSIAF 776
DB 724 LEHCYRSMVSLGLVHITNMAFLCFLGLTFVQSQPGKRNRRGLTFPMLAFTWVSF 783
OY 777 FTASVYDGKYLPPANMMAGLSSSGFGCYLPKCYVILCRPDLNSTEHF 827
DB 784 VPLLANVOAYOPAVOMGAILVICALGILVTFLPKCYVILMLPKLNTQEFF 834

RESULT 12
OY 925D9 PRELIMINARY: PRT: 858 AA.
AC 0925D9:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE SWEET TASTE RECEPTOR FAMILY 1 MEMBER 3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=2125287; PubMed=11326277;
RA Max M., Shanker Y.G., Huang L., Rong M., Campagne F., Liu Z.,
RA Weinstein H., Damak S., Margolske R.F.;
RT "Tas1r3, encoding a new candidate taste receptor, is allelic to the
RT sweet responsiveness locus Sac."
RL Nat. Genet. 28:58-63(2001).
DR EMBL; AF368024; AAK5536.1;
KW Receptor.
SQ SEQUENCE 858 AA; 94633 MW; 4E416824FDB478F8 CRC64;

Query Match 25.5%; Score 1147; DB 11; Length 858;
Best Local Similarity 32.6%; Pred. No. 6,8e-88;
Matches 271; Conservative 150; Mismatches 356; Indels 54; Gaps 16;

OY 29 SPDFLLPGDYLLAGLPLHS---GCLQVRHREPVYLCDSKSCFNEGHYLPQAMRLGVEE 85
DB 26 SOQFAOGDYIIGLPLPGSTEBATLNQAPNSTLCNR--FSPGLFLAMAKMAVEE 82
OY 86 INNSTALLPNTLYGYOLYVCSDS-ANYATRLVLSLPGOHHELOGDILHSPYLAIV 144
DB 83 INNGSALLPGLRLGYDLFTCESEPVYTKSSLMFLAKVSOSIAVCNTQOPRLAVI 142
OY 145 GPDSTNRATTAALSPFLVPMISYAASSETLSVKRQYPSFLRTIPNDKYQVETWVLLQ 204
DB 143 GPHSSELALITGKFFSFLMPQVYSASMDRLSDRETFPSFRTVPSDRVQLOAVYTLQ 202
OY 205 KFGWTWISLVGSDDYGGQLGVALENOATGOCIAFKDIMPESAQVGE--RMQCLMRH 262
DB 203 NFSMNVVAALGSDDDYGRGGLSTFSSLANARGICTAHGELVQHDTSGQOLGKVDLVLRQ 262
OY 263 LAQAGATVVVVFSSRQLARVFESVVLNLTGKVVVASEAMALSRHITGVPGIORIGVL 322
DB 263 VNOSKVVVVFLASARAVYSLEFSYSIHGSLSPKVVVASESWLTSVLVMTLPINARVGYL 322
OY 323 GVAIQRAVPGIKAEFEAYARADKEAPRCHGSCSS-----NOLCRECOAF 370
DB 323 GLQGRALLPERSHYETHAL--AADP---AFCASLMAELDEEHVWGRCPCDDI 375
OY 371 MAHTMPKILKAFMSA-----YNAYRAVYVAHGLHQLGASGACS-GRGYVPQOLLE 423

```

```

Db 376 MGNLSSGLLONLSAGQLHGHQIFATVAAYVSAQALHNHTLQCNVSHCHVSEHVLPMQDLE 435
QY 424 QIKRVHFLHKDYAFNDNDPLSSYNIITAMDNGPKMTFTVLGSSWSPVOLINNETKI 483
Db 436 NMNMSFHARDLTLQFAEGNVMEYDLKMMWQSPFVLTHTVGTFTNGT---LQLOOSKM 492
QY 484 OMHGKDNQPKSVSCSDCLSGHQRVYTGPHHCCEPCVPCGAGTFLNKSDDLYRCOPCGKEE 543
Db 493 YWPG--NQVPVSCSRCKGQVRKVGKGFHSCCTCDVCDAKGRKRPDDFTCTPCNODQ 550
QY 544 WABEGSOTCEPRRTVETALREHNSWVLLAANTLLLLLTAGLFAHMLDTPVYRSAGG- 602
Db 551 WSPKSTACLPRRPKFLAMGEPVYVLLCLVLTGLALALGSLVHMDSPVLQASGGS 610
QY 603 RLCE--LMLGSLAAGSGSLYGFGEPTRPACILRQALFGLTFLSLCYVRSQQLITF 660
Db 611 QFCGGLICLGFCL---SVLLFPGRPSSASCLAQOPAMHRLPTECLSTFLQAAETV-- 665
QY 661 KFSIKVPTFYHAWQNHGAGL---VMISSAQLLICTLWLVWTPLPAREYORFPHLY 716
Db 666 --ESELPLSMANMLCSYLRGLAMLVLLATFVEAALCAWTLIAFPPEVYTDWVLPREV 723
QY 717 MLECTENSLGFIATFLYNGLLISAFACSYLGKDLPENYNEACVYTSLENFVSWIAF 776
Db 724 LEHCYHWSVSLGLVHTTNAMLAFLCFTGLTVOSQPGRYNARGLTFAMLAFTWVSF 783
QY 777 FTTHSVYDGYLPANMAGLSSSGFGYFLPKCYVILCRPDNSTEHF 827
Db 784 VPLIANVOVAIQPVOGMGAILVCALGILVTHLPKCYVLLMPLKINTQEFF 834

```

RESULT 13

```

Q01VA4 PRELIMINARY; PRT; 858 AA.
ID 091VA4;
AC 091VA4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE PUTATIVE TASTE RECEPTOR (CANDIDATE TASTE RECEPTOR T1R3) (PUTATIVE
DE SWEET TASTE RECEPTOR TYPE 1 MEMBER 3).
GN T1R3 OR TAS1R3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A..
RC STRAIN=C57BL/6N; TISSUE-BRAIN;
RA Kitagawa M.;
RT "a putative taste receptor."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A..
RC STRAIN=C57BL/6J;
RX MEDLINE=21219400; Pubmed=11319557;
RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.;
RT "A candidate taste receptor gene near a sweet taste locus."
RL Nat. Neurosci. 4:492-498(2001).
RN [3]
RP SEQUENCE FROM N.A..
RC STRAIN=C57BL/6J;
RA Bachmanov A.A., Li X., Reed D.R., Ohmen J.D., Li S., Chen Z.,
RA Torodoff M.G., deJong P., Wu C., West D.B., Chatterjee A., Ross D.A.,
RA Beauchamp G.K.;
RT "Positional cloning of the mouse saccharin preference (Sac) locus."
RL Chem. Senses 0:0-0(2001).
RN [4]
RP SEQUENCE FROM N.A..
RC TISSUE=CIRCUMVALLATE PAPILLAE;
RX MEDLINE=21222875; Pubmed=11322794;
RA Kitagawa M., Kusakabe Y., Miura H., Ninoiwa Y., Hino A.;
RT "Molecular Genetic Identification of a Candidate Receptor Gene for

```

```

RT Sweet Taste."
RL Biochem. Biophys. Res. Commun. 283:236-242(2001).
DR EMBL; AB055708; BAB62952.1; -
DR EMBL; AF337039; AAK39436.1; -
DR EMBL; AF311386; AAL08425.1; -
DR EMBL; AB049994; BAB47181.1; -
KW Receptor.
SO SEQUENCE 858 AA; 94561 MW; 4E4168279DB478F8 CRC64;

```

Query Match 25.5%; Score 1147; DB 11; Length 858;
 Best local similarity 32.6%; Pred. No. 6.8e-88;
 Matches 271; Conservative 150; Mismatches 356; Indels 54; Gaps 16;

```

QY 29 SPDTFLPDYLLAGLFLPHS---GCLQVRHREYTLCDRSCSFNKHGHLQAMLGVEE 85
Db 26 SQOFRAGQDYLLGLFPLGSTEATLNTORTPNSTPCNR---FSPGLFLMAMKMAVEE 82
QY 86 INNSTALPLNTLLGQYLDVCSDS--ANYATLRYLSLPGOHNIELOGDLHYSPVLAVI 144
Db 83 INNGSALLPGLRGLDGLDFDTCEPYVTMKSSLMFLAKVSGSIAAYCVYQYQRPVLAVI 142
QY 145 GPDSTNRATTAALISPLVPMISTYASSETLSVAKQYPSLRTIPNDKYQVETWYLLQ 204
Db 143 GPHSESLALITGKEFSEFLMPQVSYASMDRLSDRETPSPFRTYPSDROVLQAVYTLQ 202
QY 205 KFGWTWISLVGSSDDYDGLQVALENOATGGICIAFDIMPEFSQVQDE--RMQCLMRH 262
Db 203 NFSNMWVAALSDDDYDGEGLSTPSSLANARGICIAHGGLVPQHTSQQLAKVLDVLRQ 262
QY 263 LAQAGATVVVVFSSROLARVEFESVYVLTNLGKVVAVASEANALSRHTGVQIORIGVL 322
Db 263 VNQSKVQVYVLFASARAVYSLFYSYIHHGLSPKVVAVASEMSTSDLVMTLPRIARVGYVL 322
QY 323 GVALQKRAVPEGLKAFEEYARADKEAPRPCHKSGKSS-----NQLCEQCAF 370
Db 323 GFLQRGALLPEFHSYVETHTAL--AADP---AFCAISLNELEDEHYMGQRCPCDDI 375
QY 371 MAHMPKIKAFMSMSA-----YNAVYRAVYAVAGLHOLLGASGACS--RGVYWMOLLE 423
Db 376 MGNLSSGLLONLSAGQLHGHQIFATVAAYVSAQALHNHTLQCNVSHCHVSEHVLPMQDLE 435
QY 424 QIKRVHFLHKDYAFNDNDPLSSYNIITAMDNGPKMTFTVLGSSWSPVOLINNETKI 483
Db 436 NMNMSFHARDLTLQFAEGNVMEYDLKMMWQSPFVLTHTVGTFTNGT---LQLOOSKM 492
QY 484 OMHGKDNQPKSVSCSDCLSGHQRVYTGPHHCCEPCVPCGAGTFLNKSDDLYRCOPCGKEE 543
Db 493 YWPG--NQVPVSCSRCKGQVRKVGKGFHSCCTCDVCDAKGRKRPDDFTCTPCNODQ 550
QY 544 WABEGSOTCEPRRTVETALREHNSWVLLAANTLLLLLTAGLFAHMLDTPVYRSAGG- 602
Db 551 WSPKSTACLPRRPKFLAMGEPVYVLLCLVLTGLALALGSLVHMDSPVLQASGGS 610
QY 603 RLCE--LMLGSLAAGSGSLYGFGEPTRPACILRQALFGLTFLSLCYVRSQQLITF 660
Db 611 QFCGGLICLGFCL---SVLLFPGRPSSASCLAQOPAMHRLPTECLSTFLQAAETV-- 665
QY 661 KFSIKVPTFYHAWQNHGAGL---VMISSAQLLICTLWLVWTPLPAREYORFPHLY 716
Db 666 --ESELPLSMANMLCSYLRGLAMLVLLATFVEAALCAWTLIAFPPEVYTDWVLPREV 723
QY 717 MLECTENSLGFIATFLYNGLLISAFACSYLGKDLPENYNEACVYTSLENFVSWIAF 776
Db 724 LEHCYHWSVSLGLVHTTNAMLAFLCFTGLTVOSQPGRYNARGLTFAMLAFTWVSF 783
QY 777 FTTHSVYDGYLPANMAGLSSSGFGYFLPKCYVILCRPDNSTEHF 827
Db 784 VPLIANVOVAIQPVOGMGAILVCALGILVTHLPKCYVLLMPLKINTQEFF 834

```

RESULT 14

Q923K0

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 10:18:37 ; Search time 25.2 Seconds

(without alignments)
1292.188 Million cell updates/sec

Title: US-09-819-946-2

Perfect score: 4493

Sequence: 1 MLCTARLVGLQLISCCWA.....NSTEHFQASIQDYTRRCGST 841

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1142	25.4	1078	1 CASR_HUMAN	P1180 homo sapien
2	1126.5	25.1	1085	1 CASR_BOVIN	P35384 bos taurus
3	1123	25.0	1079	1 CASR_RAT	P48442 rattus norv
4	1105.5	24.6	1079	1 CASR_MOUSE	O9QY96 mus musculu
5	741.5	16.5	1212	1 MGR5_HUMAN	P41594 homo sapien
6	727	16.2	1199	1 MGR1_RAT	P23383 rattus norv
7	727	16.2	1203	1 MGR5_RAT	P31424 rattus norv
8	719	16.0	1194	1 MGR1_HUMAN	Q13255 homo sapien
9	708.5	15.8	912	1 MGR4_RAT	Q14833 rattus norv
10	704.5	15.7	912	1 MGR2_HUMAN	P31442 rattus norv
11	692.5	15.4	872	1 MGR2_RAT	P31442 rattus norv
12	688	15.3	879	1 MGR3_RAT	P31442 rattus norv
13	684	15.2	877	1 MGR3_HUMAN	Q14832 homo sapien
14	666.5	14.8	872	1 MGR2_HUMAN	Q14416 homo sapien
15	666.5	14.8	877	1 MGR6_HUMAN	P35349 rattus norv
16	664	14.6	871	1 MGR6_RAT	P70579 rattus norv
17	652	14.6	908	1 MGR8_RAT	P47743 mus musculu
18	652	14.5	908	1 MGR8_MOUSE	O00222 homo sapien
19	650	14.5	908	1 MGR8_HUMAN	P1685 diosophila
20	626.5	13.9	976	1 MGR_DROME	Q14831 homo sapien
21	619.5	13.8	915	1 MGR7_HUMAN	P35400 rattus norv
22	619.5	13.8	915	1 MGR7_RAT	O09650 caenorhabd
23	590.5	13.1	999	1 MGR1_MOUSE	Q09650 caenorhabd
24	186.5	4.3	960	1 GBR1_MOUSE	Q09650 caenorhabd
25	186.5	4.2	940	1 GBR2_RAT	O88871 rattus norv
26	185	4.1	961	1 GBR1_HUMAN	Q9ub55 homo sapien
27	183.5	4.1	991	1 GBR1_RAT	Q92044 rattus norv
28	172.5	3.8	941	1 GBR2_HUMAN	O95899 homo sapien
29	163	3.6	938	1 NMZ1_HUMAN	O05566 homo sapien
30	158	3.5	938	1 NMZ1_RAT	P35438 rattus norv
31	156	3.5	938	1 NMZ1_MOUSE	P35438 rattus norv
32	144	3.2	1103	1 CYGF_BOVIN	O02740 bos taurus
33	135	3.0	1108	1 CYGF_HUMAN	P51841 homo sapien

34	124	2.8	986	1 CYGR_ARBP	P11528 arabacia pun
35	112.5	2.5	1125	1 CYGS_STRPU	P16065 strongyloce
36	110.5	2.5	908	1 GIK2_HUMAN	O13002 homo sapien
37	109	2.4	908	1 GIK2_RAT	P42260 rattus norv
38	108.5	2.4	540	1 ANPC_HUMAN	P17342 homo sapien
39	108.5	2.4	1108	1 CYGF_RAT	P51842 rattus norv
40	108	2.4	1350	1 XFIN_XENLA	P08045 xenopus lae
41	107.5	2.4	803	1 Z226_HUMAN	O9ayt6 homo sapien
42	107	2.4	889	1 GIK2_MOUSE	P39087 mus musculu
43	106.5	2.4	545	1 NUS5_ALBRO	P48918 albinaria c
44	106.5	2.4	557	1 GHT4_SCHPO	O59932 schizosacch
45	106	2.4	487	1 BCH2_RHOGE	Q9jpb9 rhodocyclus

ALIGNMENTS

RESULT	1	STANDARD	PRT: 1078 AA.
ID	CASR_HUMAN		
AC	P41180; Q13912; Q16379; Q16108; Q16109; Q16110;		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Extracellular calcium-sensing receptor precursor (Casr) (Parathyroid		
DE	Cell calcium-sensing receptor).		
CN	CASR OR GPRC2A OR PCARL.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_Taxid=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Pearce S.H.S., Thakker R.V.;		
RL	Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Parathyroid;		
RA	MEDLINE=95279439; PubMed=7759551;		
RT	Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M.,		
RT	Hebert S.C., Nemeth E.F., Fuller F.;		
RT	"Molecular cloning and functional expression of human parathyroid		
RL	calcium receptor cDNAs.";		
RN	J. Biol. Chem. 270:12919-12925(1995).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Kidney;		
RA	MEDLINE=95408281; PubMed=7677761;		
RT	Aida K., Koishi S., Tawata M., Onaya T.;		
RT	"Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from		
RL	human kidney.";		
RN	Biochem. Biophys. Res. Commun. 214:524-529(1995).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=96343808; PubMed=8756555;		
RT	Freichel M., Zink-Lorenz A., Hollroschl A., Hafner M., Flockert V.,		
RT	Rave F.;		
RT	"Expression of a calcium-sensing receptor in a human medullary		
RL	thyroid carcinoma cell line and its contribution to calcitonin		
RN	secretion.";		
RN	Endocrinology 137:3842-3848(1996).		
RN	[5]		
RP	SEQUENCE OF 643-908 FROM N.A.		
RA	MEDLINE=96193893; PubMed=8613532;		
RT	Bikle D.D., Ratnam A., Mauro T., Harris J., Pillal S.;		
RT	"Changes in calcium responsiveness and handling during keratinocyte		
RL	differentiation. Potential role of the calcium receptor.";		
RN	J. Clin. Invest. 97:1085-1093(1996).		
RN	[6]		
RP	VARIANTS FHH GLN-165; LYS-297 AND TRP-795.		
RA	MEDLINE=94094324; PubMed=7916660;		
RT	Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J.,		
RA	Steinmann B., Levi T., Seidman C.E., Seidman J.G.;		

"Mutations in the human Ca(2+)-sensing receptor gene cause familial hypocalcemic hypercalcaemia and neonatal severe hypoparathyroidism.";
 RL Cell 75:1297-1303(1993).
 [7]
 RP VARIANT ADH ALA-127.
 RX MEDLINE-95179179; PubMed-7874174;
 RA Pollak M.R., Brown E.M., Estep H.L., McLaine P.N., Kifor O., Park J., Hebert S.C., Seidman C.E., Seidman J.G.;
 RT "Autosomal dominant hypocalcaemia caused by a Ca(2+)-sensing receptor gene mutation.";
 RL Nat. Genet. 8:303-307(1994).
 [8]
 RP VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
 RX MEDLINE-95243222; PubMed-7726161;
 RA Chou Y.-H.W., Pollak M.R., Brandl M.L., Toss G., Argyryst H., Atkinson A.B., Papapoulos S.E., Marx S., Brown E.M., Seidman J.G., Seidman C.E.;
 RT "Mutations in the human Ca(2+)-sensing receptor gene that cause familial hypocalcemic hypercalcaemia.";
 RL Am. J. Hum. Genet. 56:1075-1079(1995).
 [9]
 RP SEQUENCE OF 1-61 FROM N.A., AND VARIANT FHH ALA-39.
 RX MEDLINE-95403641; PubMed-7673400;
 RT "Familial hypocalcemic hypercalcaemia associated with mutation in the human Ca(2+)-sensing receptor gene.";
 RL J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
 [10]
 RP VARIANTS NSHPT LEU-227 AND TYR-598.
 RX MEDLINE-9692293; PubMed-8675635;
 RA Pearce S.H.S., Trump D., Wooding C., Besser G.M., Chew S.L., Grant D.B., Heath D.A., Hughes I.A., Paterson C.R., Whyte M.P., Thakker R.V.;
 RT "Calcium-sensing receptor mutations in familial benign hypercalcaemia and neonatal hyperparathyroidism.";
 RL J. Clin. Invest. 96:2683-2692(1995).
 [11]
 RP VARIANTS ADHP THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.
 RX MEDLINE-96311554; PubMed-8733126;
 RA Baron J., Winer K.K., Yanovski J.A., Cunningham A.W., Laue L., Zimmerman D., Cutler G.B. Jr.;
 RT "Mutations in the Ca(2+)-sensing receptor gene cause autosomal dominant and sporadic hypoparathyroidism.";
 RL Hum. Mol. Genet. 5:601-606(1996).
 [12]
 RP VARIANT FHH ARG-174.
 RX MEDLINE-97442275; PubMed-9298824;
 RA Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laine N.G., Pullan P.T., Ratajczak T.;
 RT "A novel mutation (H174R) in the Ca2+-sensing receptor gene associated with familial hypocalcemic hypercalcaemia.";
 RL Hum. Mutat. 10:233-235(1997).
 [13]
 RP VARIANT FHH GLU-557.
 RX MEDLINE-21603857; PubMed-11762699;
 RA Nakayama T., Minato M., Nakagawa M., Soma M., Tobe H., Aoi N., Kosuge K., Sato M., Ozawa Y., Kamatsue K., Kokubun S.;
 RT "A novel mutation in Ca2+-sensing receptor gene in familial hypocalcemic hypercalcaemia.";
 RL Endocrine 15:277-282(2001).
 -I- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 -I- SUBCELLULAR LOCATION: Integral membrane protein.
 -I- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.
 -I- TISSUE SPECIFICITY: FOUND IN KIDNEY, BUT NOT IN BRAIN, LUNG, LIVER, HEART, SKELETAL MUSCLE, OR PLACENTA.
 -I- DISEASE: DEFECTS IN CASR ARE A CAUSE OF FAMILIAL HYPOCALCAEMIC HYPERCALCAEMIA (FHH) AND NEONATAL SEVERE HYPERPARATHYROIDISM (NSHPT), TWO INHERITED CONDITIONS CHARACTERIZED BY ALTERED CALCIUM

HOMEOSTASIS. THE MUTATIONS REDUCE THE ACTIVITY OF THE RECEPTOR. FHH AFFECTED INDIVIDUALS EXHIBIT MILD OR MODERATE HYPERCALCAEMIA, RELATIVE HYPOCALCAEMIA, AND INAPPROPRIATELY NORMAL PTH LEVELS. IN CONTRAST, NSHPT IS A RARE AUTOSOMAL RECESSIVE LIFE-THREATENING DISORDER CHARACTERIZED BY VERY HIGH SERUM CALCIUM CONCENTRATIONS, SKELETAL DEMINERALIZATION, AND PARATHYROID HYPERPLASIA. IN SOME INSTANCES NSHPT HAS BEEN DEMONSTRATED TO BE THE HOMOZYGOUS FORM OF FHH.
 -I- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT HYPOCALCAEMIA (ADH) IN WHICH THE RECEPTOR IS ACTIVATED AT SUBNORMAL CA(2+) LEVELS.
 -I- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT HYPOPARATHYROIDISM (ADHP). ADHP IS CHARACTERIZED BY HYPOCALCAEMIA AND HYPERPHOSPHATEMIA DUE TO INADEQUATE SECRETION OF PARATHYROID HORMONE. SYMPTOMS ARE SEIZURES, TETANY AND CRAMPS.
 -I- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

 CC EMBL: X81086; CAA56990.1; -
 CC EMBL: U20759; AAA86503.1; -
 CC EMBL: U20760; AAA86504.1; -
 CC EMBL: D50855; BAA09453.1; -
 CC EMBL: S83176; AAB46873.1; -
 CC EMBL: S79217; AAB35262.2; -
 CC EMBL: S68032; AAB29413.2; ALT. SEQ.
 CC EMBL: S68033; AAB29414.1; -
 CC EMBL: S68036; AAB29415.1; -
 CC EMBL: S81755; AAD14370.1; -
 CC GCRDB: GCR_1337; -
 CC GCRDB: GCR_1874; -
 CC GCRDB: GCR_2012; -
 CC GCRDB: GCR_2013; -
 CC GCRDB: GCR_2696; -
 CC GCRDB: GCR_2697; -
 CC MIM: 601199; -
 CC MIM: 145980; -
 CC MIM: 239200; -
 CC MIM: 601198; -
 CC InterPro: IPR001828; ANF_receptor.
 CC InterPro: IPR00337; GPCR_Mgr.
 CC Pfam: PF00003; 7tm_3; 1.
 CC Pfam: PF01094; ANF_receptor; 1.
 CC PRINTS: PR00248; GPCR_MGR.
 CC PROSITE: PS00979; G_PROTEIN_REC_F3_1; 1.
 CC PROSITE: PS00980; G_PROTEIN_REC_F3_2; 1.
 CC PROSITE: PS00981; G_PROTEIN_REC_F3_3; 1.
 CC PROSITE: PS00259; G_PROTEIN_REC_F3_4; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Disease mutation; Alternative splicing; Polymorphism.
 CC SIGNAL: 1 19
 CC CHAIN: 20 1078
 CC DOMAIN: 20 612
 CC TRANSFEM: 613 635
 CC TRANSFEM: 636 649
 CC TRANSFEM: 650 670
 CC TRANSFEM: 671 681
 CC TRANSFEM: 682 700
 CC TRANSFEM: 701 724
 CC TRANSFEM: 725 745
 CC TRANSFEM: 746 769
 CC TRANSFEM: 770 792
 CC TRANSFEM: 793 805
 CC TRANSFEM: 806 828
 CC TRANSFEM: 829 836
 CC TRANSFEM: 837 862
 CC TRANSFEM: 863 1078
 CC DOMAIN: 863 1078
 CC I (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC II (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC III (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC IV (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC V (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).
 CC VI (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC VII (POTENTIAL).
 CC CYTOPLASMIC (POTENTIAL).

Query Match 25.1%; Score 1126.5; DB 1; Length 1085;
 Best Local Similarity 30.7%; Pred. No. 2.3e-74;
 Matches 274; Conservative 163; Mismatches 362; Indels 93; Gaps 22;

14 LISCCAFKCHST-----ESSPD--FLPBDYLLAGLFPPLHSGCL-----QVRHREPTLTD 63
 3 LKSCCMILLAFSTWCTSAVGPDOAKKGDIIILGIFPIHFGVAVKDOLKSPESVECI 62
 64 RSCSENEHGHFQAMRLCGLVEEINNSTALLPNTLLGQYDYVSDSAGVATLRLVSLG 123
 63 R---YNERGFRLLQAMIFAEIENSPALLPMNTGLRIFDNCNYSK--ALEATISFPA 117
 124 QHHIEIÖG-----DLHSPVLAIVGPDSTNRATTAALLSPFLVPMIStAAStETLSV 178
 118 OKKIDSLINDEFNCSEHPISTIAVVGATSGSTFVAVNMLGLEIFYIPQVSASSSLLEN 177
 179 KQOYSEFLRTINDKYQVETWVLLQKPGWTMISLVGSDDYQGLGVQALLENQATGQIC 238
 178 KQOFSFLRTINDKYQVETWVLLQKPGWTMISLVGSDDYQGLGVQALLENQATGQIC 237
 239 IAFKIDMPSPAGVDEEMOCLMRHLAQAGATVYVSSRQLARVPESVVLNLTGKVVW 298
 238 IDFSELI--SQYSEKIQVVEVIONSTAKIYVFFSSGPDLEPLIKETVRNRITGRIMW 295
 299 ASEANALSHRITGVGRIGRIGVLAIOKRAVPGIKAF-----EAYARADKEA 348
 296 ASEANALSHRITGVGRIGRIGVLAIOKRAVPGIKAF-----EAYARADKEA 348
 349 PRPCH-----KSCWSSNO-----LCRECAFMAMHTPKAKESM 383
 356 TENCHLQEGAKGPLPDPFLRGHEEGARLSNPTAFRPLCTGEEINISSEVERPYMDYTHL 415
 384 SSAYANRYAVRYAVAGLHOLLG-----ASGASR--GRVYPMOILEIHHVFLHLK- 434
 416 RISTYNYAVLYAIAHALQDITYICIGRGFLTNGSCADIKKVERMOVLKRLHNTFSNNG 475
 435 DTVAENDNBDPLSSYNIIAMDNNGPRMPTVLSSSTWSPV-----OLINNETKI 483
 476 EGVTEDECGDLANGNYSIIWMHLSPED-----GSIFYKRVGYVNYAKKGERLFINDEKI 529
 484 OMHGKDNQPKVCSDDCLEGHOR--VYGFHHCCEPCVCGAGTFLNKSDDLRCOPGKE 542
 530 IMSGFSREVPFNSCSDCLAGTRKGIIECEPTCECEVCPCGEXSDEEDASACDKCPCPD 589
 543 EVAPRGSOTCPRTVYETALREHTSMVLAANTLLILLGAGLFAWHLDPVYVSAG 602
 590 FMSNEHNTSCIAKEIEFLSWTEPPGIALFLFAYLVGFATKFRNTPYVATNR 649
 603 RLCLMTLGLSAGSGLYGFEGEPTRPACILRQALFALGFTIFLSCLTVRSFOLIIIFK 662
 650 ELSTYLLFSLLCCFSSLEFFIGEPQDMWCRLRQPAFGISFVLCISILVTKRVLVFE- 708
 663 SKRVLPFH--ANVQNHGAGLEFYMISAOGLICTLWLVVWTPLPAREYORFPLVMECT 721
 709 -AKIPTSFRRKMGWNTLFLVFLCTFMQIVCAIWLNPAPSSYXNHLEDEIIIFITCH 767
 722 ENN--SIGFTILFLYNGLSISAFACSYLGDLPENYKACVTSILNPFVSMIAFFT 779
 768 EGSIALAGFLIG--YTCLIAAICFFAFRSKRLPENFNAKTITISMLFFFTWISFTIRA 825
 780 -ASVYDGLKLPANMAGLSLSGFGYFLPRCVYILCRPLDUNSTEHFQAS 830
 826 VASTY--GKRVSAVEYIALAASFGLLACIFFNKVVYITLKKPRNTIEEVRC 876

RESULT 3
 CASR_RAT STANDARD: PRT: 1079 AA.
 AC P48442:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CASR) (Parathyroid
 DE Cell calcium-sensing receptor).
 GN CASR OR GPCR2A OR PCAR1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney outer medulla;
 RX MEDLINE=95116508; PubMed=7816802;
 RA Ricciardi D., Park J., Lee W., Gamba G., Brown E.M., Hebert S.C.;
 RT "Cloning and functional expression of a rat kidney extracellular
 RT calcium/polyvalent cation-sensing receptor".
 RL Proc. Natl. Acad. Sci. U.S.A. 92:131-135(1995).
 RP SEQUENCE OF 1-294 FROM N.A.
 RC STRAIN=MISTAR.
 RX MEDLINE=95241465; PubMed=7724534;
 RA Ruat M., Snowman A.M., Snyder S.H.;
 RT "Calcium sensing receptor: molecular cloning in rat and localization
 RT to nerve terminals".
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3161-3165(1995).
 CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: U10354; AAC52149.1; -
 CC EMBL: U20289; AAC52195.1; -
 CC GCRDB: GCR_1449; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000377; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRGR.
 DR PROSITE: PS00979; G_PROTEIN_RECPE_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECPE_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECPE_F3_3; 1.
 DR PROSITE: PS00982; G_PROTEIN_RECPE_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1
 FT CHAIN 20 1079
 FT DOMAIN 20 612
 FT TRANSMEM 613 635
 FT DOMAIN 636 649
 FT TRANSMEM 650 670
 FT DOMAIN 671 681
 FT TRANSMEM 682 700
 FT DOMAIN 701 724
 FT TRANSMEM 725 745
 FT DOMAIN 746 769
 FT TRANSMEM 770 792
 FT DOMAIN 793 805
 FT TRANSMEM 806 828
 FT DOMAIN 829 836
 FT TRANSMEM 837 862
 FT DOMAIN 863 1079
 FT CARBOHYD 90 130
 FT CARBOHYD 130 130
 FT CARBOHYD 261 261
 FT CARBOHYD 287 287
 EXTRACELLULAR CALCIUM-SENSING RECEPTOR.
 I (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 II (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 III (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 IV (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 V (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 VI (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 VII (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL)
 SQ SEQUENCE 1079 AA; 120867 MW; D7664550361F9736 CRC64;

Query Match 25.0%; Score 1123; DB 1; Length 1079;
 Best Local Similarity 30.9%; Pred. No. 4,1e-74;
 Matches 275; Conservative 162; Mismatches 360; Indels 92; Gaps 22;

16 SCC---MAFACHSTRESSPD--FTLPGLDLAGLPLHSGCL---QVHREPVTLCDSC 66
 Db 5 SCCALLALLAHSSAYGPDRQKKGDIILGLFPIHGVAKKODDLSPRESEICIR-- 62
 QY 67 SFNEGHYHLPQAMRLGVEINNSTALPNTLGYOLVDCSDSANVATLRLVSLPGQH 126
 Db 63 -YNFRGFRMLQAMTFAIEEINSSPLPNTLGYRIPTCTMYSK--ALEATLSFVAQNK 119
 QY 127 IELDG-----DLHSPVYAVIGPDSTNRRAATTAALLSPFLVPMISTAASSETLSVRQ 181
 Db 120 IDSLNIDFCNCSEHPISTIAVVGATGSGVSTAVANLGLFYIPQVSYASSRLSNKQ 179
 QY 182 YPSFLRTIPNDKYQVETWLLQKFGMTWISLVGSSDDYGGOLVQALENOATGQICIAF 241
 Db 180 YKSEFLRTIPNDHQATMADIETFRMVMVGTIAADDDYGRGICKEFEAEFERICIDF 239
 QY 242 KDIMPFSQVDERMOCLMRHLAAGATVYVVFSSROLARVFESVLTNLTKGVAVASE 301
 Db 240 SELI--SQYSEDEEIQVVEYQNSTAKIYIVFSSGPDLEPLIKETVRNITGRIMLASE 297
 QY 302 AMALSRHTITGPGIORIGMVGVALQKRAVPGIKAF-----EAAVRAADKEARP 351
 Db 298 AMASSSLIAMPDEYFHVGGTIGFGIKAGQIPGFEFLQKVPKRKSVHNGFAEFMEETFN 357
 QY 352 CH-----KG-----SMCSSNOLRECOAFMAHMKPKAKARMSMA 386
 Db 358 CHLGGAGPRLPVDTFVNSHEGGRRLNLSSTAFRLPCTGSDENISVETPYMDYHLLAIS 417
 QY 367 YNAVRAVVAHGLHQLIGC-----ASGACSR--GRAYPMOLLEQIHKVHFLHK-DTV 437
 Db 418 YNVYLAIVYIAHALODITYCLPGRCLFTNGSCADIKKVEAMOVLEKHLRLHNTNMGRQV 477
 QY 438 AFNQRDRLSSYNTIATMMNGRKPTFTYLGSTWSPV-----QLNINERKIOMH 486
 Db 478 TFDECGDLVGNSTIINMHLSPED-----GSIVFEVEGYVYVYAKKGERLFTNEKILMS 531
 QY 487 GKDNQVPRKSVSCDLECHQR--VVTGFHHCCFECVPCGAGTFELNKSRLDYRCOPCKEEMA 545
 Db 532 GFSRREVPFNSCRDQATRKGIIEGPTCECEPCDEGSEGTDSACDKCDDRWIS 591
 QY 546 PEGSOTCEPRVVFALREHTSWVLAANTLLLLLLCTAGLFAVHLDTPVVSAGRLC 605
 Db 592 NENHTSIAKEIEFLAMEPFGIALTLFAVLGIFLTAFLVGFIFRMTPIYKATNRELS 651
 QY 606 FLMGSLAAGSGSLGEGEPTRPACLRQALFALGFTIIFLSCLVYRSFQILIIKFKSTK 665
 Db 652 YLLPSLCLCFSSSLFEGEPQDMTCRLQPAFGISFVLCISCLIVKTNRLVLEVE--AK 709
 QY 666 VPTFYH-AMVONHGAGLFVMISSAQLICTLWLVVMTPLPAREYQRPPLHMLCETFN 724
 Db 710 IPTSPHRKRWGINTQFLVLCFPMQILICIMLTAPSSRNHLEDEIIFICHGSS 769
 QY 725 --SLGFIATLYNGLSLTSAFACSTYGLKDLPEVNEAKCVTFSLFNFVSWIAFTT--AS 781
 Db 770 LMAISLIG--YTCLLAICEFFFAKSRKLPENNEAKIFTESMILFIYVWISFIPAYAS 827
 QY 782 VYDGKYLPAAMMAGLSSLSGFGYFLPKCYVILCRDNLNSTERFQAS 830
 Db 828 TY-GKFSVAVEVIALIASFGILACIFPNKYIILFKPSRMTIEVRSS 875

RESULT 4
 ID CASR_MOUSE STANDARD; PRT: 1079 AA.
 AC Q9QY96; Q08968; Q88519; Q9QY95; Q9QZ08; Q9R1D6; Q9R1V2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CaSR) (Parathyroid
 DE Cell calcium-sensing receptor).
 GN CASR OR GPRC2A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxId=10090.
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RX MEDLINE=20092890; PubMed=10625662;
 RA Oda Y., Tu C.-L., Chang W., Crumrine D., Koemueves L., Mauro T.,
 RA Elias P.M., Bikle D.D.;
 RT "The calcium sensing receptor and its alternatively spliced form in
 RT murine epidermal differentiation.";
 RL J. Biol. Chem. 275:1183-1190(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC STRAIN=BLACK SWISS X 129/SVJ; TISSUE=Kidney;
 RX MEDLINE=20119279; PubMed=10652312;
 RA Pi M., Garner S.C., Flannery P., Spurney R.F., Quarles L.D.;
 RT "Sensing of extracellular cations in CaSR-deficient osteoblasts.
 RT Evidence for a novel cation-sensing mechanism.";
 RL J. Biol. Chem. 275:3256-3263(2000).
 RN [3]
 RP SEQUENCE OF 256-600 FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC TISSUE=Epiphyseal cartilage;
 RX MEDLINE=20043955; PubMed=10579354;
 RA Chang W., Tu C., Chen T.-H., Komuves L., Oda Y., Pratt S.A.,
 RA Miller S., Shoback D.;
 RT "Expression and signal transduction of calcium-sensing receptors in
 RT cartilage and bone".
 RL Endocrinology 140:5883-5893(1999).
 RN [4]
 RP SEQUENCE OF 507-582 FROM N.A. (ISOFORM A).
 RC STRAIN=NMRI; TISSUE=Brain;
 RA Hildbrand J., Ammon H.P.T., Wahl M.A.;
 RT Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
 RL [5]
 RP SEQUENCE OF 562-814 FROM N.A.
 RC TISSUE=Kidney;
 RA Moawad T.I., Ricciardi D.;
 RT Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 646-799 FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=97231187; PubMed=9076582;
 RA Quarles L.D., Hartle J.E. II, Siddhanti S.R., Guo R., Hinson T.K.;
 RT "A distinct cation-sensing mechanism in MC3T3-E1 osteoblasts
 RT functionally related to the calcium receptor".
 RL J. Bone Miner. Res. 12:393-402(1997).
 CC - FUNCTION: SENSES CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC - SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL AF110178; AAD28371.1; -

CC DR EMBL AF110179; AAD28372.1; -

CC DR EMBL AF128842; AAD40638.1; -

CC DR EMBL AF068900; AAC19388.1; -

CC DR EMBL AB027140; BAA77688.1; -

CC DR EMBL AF002015; AAC53252.1; -

CC DR EMBL AF159565; AAF00193.1; -

CC DR MGD:1351351; Casr

CC DR InterPro: IPR001828; ANF_receptor.

CC DR InterPro: IPR000337; GPCR_Mgr.

CC DR Pfam: PF00003; 7tm_3; 1.

CC DR Pfam: PF01094; ANF_receptor; 1.

CC DR PRINTS: PR00248; GPCRMR.

CC DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.

CC DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.

CC DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.

CC DR PROSITE: PS50259; G_PROTEIN_RECP_F3_4; 1.

CC DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;

KW Alternative splicing.

FT SIGNAL 1 19

FT CHAIN 20 1079

FT DOMAIN 20 612

FT TRANSMEM 613 635

FT DOMAIN 636 649

FT TRANSMEM 650 670

FT DOMAIN 671 681

FT TRANSMEM 682 700

FT DOMAIN 701 724

FT TRANSMEM 725 745

FT DOMAIN 746 769

FT TRANSMEM 770 792

FT DOMAIN 793 805

FT TRANSMEM 806 828

FT DOMAIN 829 836

FT TRANSMEM 837 862

FT DOMAIN 863 1079

FT TRANSMEM 90 130

FT CARBOHYD 130 261

FT CARBOHYD 261 287

FT CARBOHYD 287 366

FT CARBOHYD 366 446

FT CARBOHYD 446 468

FT CARBOHYD 468 488

FT CARBOHYD 488 541

FT CARBOHYD 541 594

FT CARBOHYD 594 537

FT VARSPLIC 461 537

FT VARSPLIC 45 45

FT CONFLICT 304 304

FT CONFLICT 410 410

FT CONFLICT 566 566

FT CONFLICT 595 595

FT CONFLICT 610 610

FT CONFLICT 814 814

FT CONFLICT 889 889

FT CONFLICT 906 909

FT CONFLICT 1057 1057

FT CONFLICT 1064 1064

FT CONFLICT 1076 1076

FT SEQUENCE 1079 AA; 120839 MM; AAF8BD472736DE CRC64;

Query Match 24.6%; Score 1105.5; DB 1; Length 1079;

Best Local Similarity 30.5%; Pred. No. 7.7e-73;

Matches 271; Conservative 165; Mismatches 359; Indels 93; Gaps 22;

QY 14 LISCWAFACHSTRESSPD--FTLPGLDYLALGLPLHSGCL---QVHRHPEVTLCDRSCS 67

DB 10 LIALITW----HSSAYGPDQRAQKKDIIILGGLPIHFHGYAAKQDLAKSRPESVCEIR--- 62

QY 68 FNEHGYLEOAMKRLGVEEINNSTALLPNTITGLYDVCSDSANVATLRLVSLPGQHII 127

DB 63 YNRCFRMLQAMFAIEEINSSPALLPNTITGLYDVCSDSANVATLRLVSLPGQHII 120

QY 128 ELQG-----DLHSTVLAIVGPDSTNKAATTAALLSFVPMISYASSSETLSVKRXY 182

DB 121 DSIUNDEFNCSEHPISTIAVVGANGSSVAVANLLGLFYIPQVYASSSRLSKNPF 180

QY 183 PSFLRTIPNDKYOVETMVLLOKFGWTYSLVGSDDYQOLGVOALENOATGOCIGIAR 242

DB 181 KSLRTIIPNDKHATMAADIIIEFRNHWGTIAADDDYGRPIEKRFEAREERICIDDS 240

QY 243 DIMPSAQAQVDERMOCLMRHLAQAATVAVVSSROLAIVFESVYLNLGKVVAVSEA 302

DB 241 ELI--SQYSDDEEILQGVNEVIONSTAKVIVFSSGDEPLLEKEIYRNITGRIMLASEA 298

QY 303 WALSRITGVPGIQRGMVLGAIQKRAVGLKAF-----EENARADKEAPRC 352

DB 299 WASSSLIAMPVEYFHVVGITIGFLKAGQIPGREFLQKVHPKRSVHNGFAKEFEETENC 358

QY 353 H-----KG-----SMCSSNOLCRECOAFMAHTMKLAFSMSAY 387

DB 359 HLDGAKGGLPDTFVRSHEBGNRLNSTAFRLCTGDENINSVETPMGEHLRIY 418

QY 388 NAYRAYVAVAHGLHQLLG-----ASGACSR-GRVYPMOLLEQIHVHPLIKR-DIVA 438

DB 419 NYLAVYSIAHALQDIYTCPLPGRGLTFNCSADIKVEAMQVLKHLRLHFTNNMEQYT 478

QY 439 FNDNRPLSSYNTIAMDNMGPMTFTVIGSSWSPV-----QINIEKTIQWNG 487

DB 479 FDCGDLVGNYSITNNHLSPED-----GSIVKEVGYNNYAKKGERLFINEGKILMSG 532

QY 488 KNOYPRKSVSCDPLEGHOR-VVTGFHCCFECVPCAGTFLNKSPLYRCQCGKEENAP 546

DB 533 FSRVEVFSKSCSDCQATKGIIEGEPCCFECVCEPDEYSGEDDASCDKCPDPFVN 592

QY 547 EGSQTCPPRTVFLALREHTSVLLAANTLLLLLGTAGLFAMHLDPVVSAGRGLCF 606

DB 593 EYNTSCIAEIEIFLAMEFEGIALTLFVAVLGFVIFKRFNPIVAKATNELSY 652

QY 607 IMGSLAAGSSLYGFGFPTPACILROALFALGTIFLSCTVRSFPLIIFKSTV 666

DB 653 LILFSLCCFSSSLFTIGEPQMTCHLRQPARGISFVLICILYKTNVLYVE--AKI 710

QY 667 PTFFVH-AMVQNHAGLFFWISSAOLILCTWLVVWTPPLPAREYORFPLVMECTFN- 724

DB 711 PTFHKKMGMLNQLFLVFLCFEMQIVICIMLYTAPPSYNNHEDEIRITCHEGSL 770

QY 725 -SLGFLAFLYGLSLISAFACSYLGLKDPENYNEAKCYTESLRFVSWIAFFT-ASV 782

DB 771 MALGSLIG--YTCLAAALCFEFAFRKSLRPENFENEKFTTFEMLIFFIWMISFIPAYAST 828

QY 783 YDGKYLPAANMAGLSLSSGFGYFLPKCYILCPDPNSTEHOAS 830

DB 829 Y-GKVSAYEVAIILASFGLLACIFPNKYIILFRPSNMTIEVRSS 875

RESULT 5

ID MGR5_HUMAN STANDARD; PRT; 1212 AA.

AC PA1594;

DT 01-NOV-1995 (rel. 32, Created)

DT 01-OCT-1996 (rel. 34, Last sequence update)

DT 16-OCT-2001 (rel. 40, Last annotation update)

DE Metabotropic glutamate receptor 5 precursor.

GN GRM5 OR GPRC1E OR MGLUR5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

[illegible]

```

RESULT 6
ID MGR1_RAT STANDARD: PRT: 1199 AA.
AC P23385:
DT 01-NOV-1991 (rel. 20, Created)
DT 01-NOV-1991 (rel. 20, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 1 precursor.
GN GRM1 OR GPRC1A OR MGLUR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=91156047; PubMed=1847995;
RA Masu M., Tanabe Y., Tsuchida K., Shigemoto R., Nakanishi S.;
RT "Sequence and expression of a metabotropic glutamate receptor."
RL Nature 349:760-765(1991).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9202526; PubMed=1656524;
RA Houamed K.M., Kujper J.L., Gilbert T.L., Haldeman B.A., O'Hara P.U.,
RA Mulvihill E.R., Almers W., Hagen F.S.;
RT "Cloning, expression, and gene structure of a G protein-coupled
RT glutamate receptor from rat brain."
RL Science 252:1318-1321(1991).
RN 13
RP ALTERNATIVE SPLICING (ISOFORM 1B).
RC TISSUE=Brain;
RX MEDLINE=92110002; PubMed=1309649;
RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
RT "A family of metabotropic glutamate receptors."
RL Neuron 8:169-179(1992).
RN 14
RP ALTERNATIVE SPLICING (ISOFORM 1C).
RC TISSUE=Brain;
RX MEDLINE=93066232; PubMed=1438218;
RA Pin J.-P., Maheux C., Prezeau L., Bockaert J., Heinemann S.F.;
RT "Alternative splicing generates metabotropic glutamate receptors
RT inducing different patterns of calcium release in Xenopus oocytes."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10331-10335(1992).
RN 15
RP FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
RN MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
RN CALCIUM SECOND MESSENGER SYSTEM. MAY PARTICIPATE IN THE CENTRAL
RN ACTION OF GLUTAMATE AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
RN THE HIPPOCAMPUS AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
RN CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
RN CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A (SHOWN HERE), 1B AND 1C;
RN CC ARE PRODUCED BY ALTERNATIVE SPLICING. 1B AND 1C ARE C-TERMINALLY
RN CC TRUNCATED FORMS OF 1A.
RN CC -1- TISSUE SPECIFICITY: IS PREDOMINANTLY EXPRESSED IN CEREBELLAR
RN CC PURKINJE CELLS, CA2-CA3 PYRAMIDAL CELLS OF THE HIPPOCAMPUS, AND
RN CC MITRAL AND TUFTED CELLS OF THE OLFACTORY BULB.
RN CC -1- MISCELLANEOUS: ACTIVATED BY QUISOALATE > GLUTAMATE > IBOVENATE >
RN CC 2-AMINO-3-PHOSPHONOPROPIONATE.
RN CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
RN CC STRONGEST, TO MGLURS.
RN CC -----
RN CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RN CC between the Swiss Institute of Bioinformatics and the EMBL outstation
RN CC the European Bioinformatics Institute. There are no restrictions on its
RN CC use by non-profit institutions as long as its content is in no way
RN CC modified and this statement is not removed. Usage by and for commercial
RN CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
RN CC or send an email to license@isb-stb.ch).
RN CC -----
RN DR EMBL; X57569; CAA40799.1; -
RN EMBL; M61099; AAA19497.1; -

```

```

DR EMBL; S48085; AAB24138.1; -
DR PIR; S15362; S15362.
DR PIR; AA1939; AA1939.
DR GCRDB; GCR_0216; -.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF01094; ANF_3; 1.
DR PRINTS; PR00248; GPCR_MGR.
DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
DR PROSITE; PS50259; G_PROTEIN_RECP_F3_4; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family; Alternative splicing.
FT CHAIN 1 18
FT SIGNAL 1 18
FT DOMAIN 19 1199
FT TRANSMEM 19 592
FT TRANSMEM 593 615
FT DOMAIN 616 629
FT TRANSMEM 630 650
FT TRANSMEM 651 661
FT TRANSMEM 662 680
FT DOMAIN 681 706
FT TRANSMEM 707 727
FT DOMAIN 728 750
FT TRANSMEM 751 772
FT TRANSMEM 773 785
FT TRANSMEM 786 808
FT DOMAIN 809 814
FT TRANSMEM 815 840
FT DOMAIN 841 1199
FT DOMAIN 1014 1034
FT DOMAIN 1074 1080
FT DOMAIN 1126 1135
FT DOMAIN 1140 1199
FT CARBOHD 98 98
FT CARBOHD 223 223
FT CARBOHD 397 397
FT CARBOHD 515 515
FT VARSPPLIC 887 906
FT VARSPPLIC 907 1199
FT VARSPPLIC 888 887
FT VARSPPLIC 898 1199
FT SEQUENCE 1199 AA; 133235 MW; EEB5A04C50694B9F CRC64;

Query Match 16.2%; Score 727; DB 1; Length 1199;
Best Local Similarity 26.7%; Pred. No. 3,6e-45;
Matches 230; Conservative 155; Mismatches 362; Indels 116; Gaps 32;

OY 34 LRGDYLAALFPLHSGCLGVRRHREY-TLCDSCS--FNEHYHFLQARLGLVEEINNST 90
DB 41 MGDVITIGLFVSH-----HQPAREKYPEKCEIRQYQIQVNEAMFHLDKINADP 93
OY 91 ALPNITLQYLYVDCSDSANVYATFLVSLPGQHIELQGLL----- 134
DB 94 VLLPNTLSEIRDSCHWS-----SVALEQSEIFRISLISIRDEKDLNCLDP 143
OY 135 -----HSPYVLAVITGDSINRAATTAALLSPLYVPISTASSETLSVKRQPSFLR 187
DB 144 GQTLPPGRTRKPIAGVIGSSSSVVALQVNLQLEPDIPIAVSATSIDSLTKYFLR 203
OY 188 TLPNDYQVEYVWLLQKGTWISLVGSSDDYGGOLGVALNOATGOCIAFKDIMP 247
DB 204 VVPSDILQARAMDYKRNMTYVSAVHTEGVSGMDAFELAAQBSCLTAHSD--KI 261
OY 248 SAQVDERMOCLMRHLAQ--AGATVAVVFSRQLARVFESVYLTNLTKGY--WVA 304
DB 262 YSNAGEKSFDRLLRKRLRELPKARVYVCCGEMTYRGLLSABRRLGYVGESFLISD 321
OY 305 LSRHITGVPGIORIGKAVNLGVALQKRAVPGLKAFPEAY--ARADKAPRCPCHGSC 362

```

```

Db 322 DRDEIEGEYEANG---GITIKLOS--PEVRSFDDYFLKLDLNTNRPMPPEFMOHRFQ 377
Qy 363 LCRCQAFMAATM--PKIKAF-----SMSSAY-----NAPVAVAAHGL---HQLL 404
Db 378 -CR-----LPGHLEENPNEKVCCTGNESEENYVODSKMGVYINAIYAAHGLONNHGLC 432
Qy 405 GCASGACSR-----GRVYPMOLLEQIKHVFHL-LHKDTAFNDNRDPLSSNIITAMDV-N 457
Db 433 PGHGLCDAMKPIDGR-----KLDFLIKSSFVSGEYWEDEKDGADGRDINMLQYTE 488
Qy 458 GPKRTFVLSGSSVSPVOLINETKIQHGMKDNVPSKSVSSDCLGHRVY-TGFHHC 516
Db 489 ANRDYVHVG--TWHEGLINIDYKIQMN-KSGMV-RSVCEPCLKGQIKVIRKGEVSC 544
Qy 517 FECPGCGAGTFLNKSIDLXRCOPCKEEMAPRGSGTCFRTVFLARHTSWVLLAATL 576
Db 545 WICPACKEHEFV--ODETCRACDLGWWPNAELTGCEPIPVYLEWSDIESIIIAFSC 602
Qy 577 LLLLLGTAGLFAHMLDTPVVSAGRLCFMLGSLAGSGSLGFECEPRPACILRQA 636
Db 603 GILVTLFVTLFVLYKRDTPVVKSSRELCTIILAGIFLGVCFTLLAKPTTSCYLOR 662
Qy 637 LFAIGFTIISCLTVRSFQIIIF----KESTVPTFYHAMVONHGAGLFVMTSS--A 688
Db 663 LVGSANMCYSALVTKTRIRIARILAGSKKIKCTKRPRMSMAQ-----VIASILIS 715
Qy 689 AQLICLTLVYVTPPLAREQRPPLVMECTETNSGLFLAPLNGLSISAFACYL 748
Db 716 VOLTLVTLIMEPPMLLSPISKEVYL--CMTSLNGVAPVPGVGLLMSCTYYAFK 773
Qy 749 GKDLPEVNEAKCYTFSLLFNFVSIAFETASVYDG-KYLPANMAGLSLSGFGY 807
Db 774 TRNPANFNEKXYIAFTMYTTCITLAF--VPYIFGSNXYIITTCRAVSLVVALGCM 830
Qy 808 FLPRCYVILCRPDINSTEHFQAS 830
Db 831 FTPKWTIIIAKPERNVSAPFTTS 853

RESULT 7
MGR5_RAT
ID MGR5_RAT STANDARD; PRT; 1203 AA.
AC P31424;
DT 01-JUL-1993 (Rel. 26, Last Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 5 precursor.
GN GRM5 OR GPRC1E OR MGLUR5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX SEQUENCE FROM N.A.
RP TISSUE=Brain;
RA Minakami R., Katsuki F., Sugiyama H.;
RA "A variant of metabotropic glutamate receptor subtype 5: an
RT evolutionally conserved insertion with no termination codon.";
RT Molecular characterization of a novel metabotropic glutamate
RT receptor mGLUR5 coupled to inositol phosphate/Ca2+ signal
RT transduction.";
RL J. Biol. Chem. 267:13361-13368(1992).
RN [2]
RP SEQUENCE OF 859-923 FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=93343913; Pubmed=7688218;
RA Minakami R., Katsuki F., Sugiyama H.;
RA "A variant of metabotropic glutamate receptor subtype 5: an
RT evolutionally conserved insertion with no termination codon.";
RT Biochem. Biophys. Res. Commun. 194:622-627(1993).
CC -I- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM AND GENERATES A CALCIUM-ACTIVATED

```

```

CC CHLORIDE CURRENT.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS, 5A (SHOWN HERE) AND 5B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER BY AN INSERTION OF
CC 32 RESIDUES.
CC -I- TISSUE SPECIFICITY: WIDELY DISTRIBUTED IN NEURONAL CELLS OF THE
CC CENTRAL NERVOUS SYSTEM.
CC -I- MISCELLANEOUS: ACTIVATED BY QUISQUALATE > GLUTAMATE > IBOTENATE >
CC TRANS-1- AMINOCYCLOPENTYL-1,3-DICARBOXYLATE.
CC -I- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR1.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D10891; NOT ANNOTATED_CDS.
CC EMBL: S64315; AB27666.1; -.
CC PIR: A42916; A42916.
CC GCRDB: GCR_0444; -.
CC GCRDB: GCR_0760; -.
CC InterPro: IPR001828; ANF_receptor.
CC InterPro: IPR000337; GPCR_Mgr.
CC Pfam: PF00003; 7tm_3; 1.
CC Pfam: PF01094; ANF_receptor; 1.
CC PRINTS: PR00248; GPCRMR.
CC PROSITE: PS00979; G-PROTEIN_RECEP_F3_1; 1.
CC PROSITE: PS00980; G-PROTEIN_RECEP_F3_2; 1.
CC PROSITE: PS00981; G-PROTEIN_RECEP_F3_3; 1.
CC PROSITE: PS50259; G-PROTEIN_RECEP_F3_4; 1.
CC K1: G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Multigene family; Alternative splicing.
CC SIGNAL 1 20
CC CHAIN 21 1203 METABOTROPIC GLUTAMATE RECEPTOR 5.
CC FT DOMAIN 22 578 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 579 601 I (POTENTIAL).
CC FT DOMAIN 602 615 CYTOPLASMIC (POTENTIAL).
CC FT TRANSSEM 616 636 II (POTENTIAL).
CC FT DOMAIN 637 647 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 648 666 III (POTENTIAL).
CC FT DOMAIN 667 692 CYTOPLASMIC (POTENTIAL).
CC FT TRANSSEM 693 713 IV (POTENTIAL).
CC FT DOMAIN 714 736 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 737 758 V (POTENTIAL).
CC FT DOMAIN 759 771 CYTOPLASMIC (POTENTIAL).
CC FT TRANSSEM 772 794 VI (POTENTIAL).
CC FT DOMAIN 795 800 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 801 826 VII (POTENTIAL).
CC FT DOMAIN 827 1203 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 377 377 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT VARSPLIC 876 907 MISSING (IN ISOFORM 5A).
CC SO SEQUENCE 1203 AA; 131885 MW; 99CA51E9E11C1EA4 CRC64;

```

```

Query Match 16.28; Score 727; DB 1; Length 1203;
Best Local Similarity 27.28; Pred. No. 3.6e-45;
Matches 235; Conservative 151; Mismatches 356; Indels 122; Gaps 33;

Qy 34 LPGYLAGLPEPLSGCLQVHRPEV-TLCDRSQSF--NENGYHFOAMRGVEINNST 90
Db 31 MPGLIILGALFSVH-----HGFYVDKVKHERKGAAREQIGQIVEMAMTLTERINS 83
Qy 91 ALPNITLGLGYLDVNCSDSANVYATLRYLSLPGHHIELQGLDLYHS----- 137

```

DB 84 TLLPNTTLCCELRSCMBSA-----VALEOSIEFIRDSLSSEEEGLVRCVDGS 133
 QY 138 -----PVLAVIGPDSRRAATTAALLSPFLVPMISVYASSSETLSVKRQYPSFLTIPN 191
 DB 134 SSFRRKRPVIGVIGRESSSVIAIOVNLLOLFNIPQIAVATSMDSKDLFKYFMVVS 193
 QY 192 DKYOVETVLLQKRGWTVISLVGSSDDYGOLGVALLENQATQGCICIA--FKDIMPESA 249
 DB 194 DAQQAARAVDIYKRYMWTYVSVAHFEGNGESGMEAFKMSAKEGICIAHSKYISNAGE 253
 QY 250 QYGDERMOLMHLNAGATVYVVSRROLARVFESVYLVLTNIGK--VWVASEANALSH 308
 DB 254 QSDPDKLKLKRLSHLPK--ARVVAECGEGMTVGLLMMARLGLAGEFLLGSDGMADRYD 311
 QY 309 ITGVPGIORIGWLVGVAIOKRAVPGLKAFEEAVARADKRAPRCHGSCWSSNOLCRECO 368
 DB 312 VV--DQYGR-EAVGGITLIGL--PDYKMWDDYLL--KLRPENLNKPM-----FQ 355
 QY 369 AFMAHTMP-KLKAFSM-SSAYN-----AVRAYVAHGLH--QL 403
 DB 356 EFWOHRFOCRLEGFQENSKYNTKNSLTLRTHVQDSKMGFVINAISYMAVGLHNMQM 415
 QY 404 LGCA--SGACSR-----GAVYPMQLLEQJHXYHFL-LHKDYAFNDRNPRLSYNIAMD 455
 DB 416 SLCPGYAGLCDAMKPIDGR-----KLDSLMKTNFTGVSGDMILFDENGSPGRYTEIMNK 471
 QY 456 WNGPKWTFVLSSSTWSPQOLINETKIQMHGKNOVPKSCSSDCEGHORVY-TGFHH 514
 DB 472 EMGKQV--FLYINVGSDNDELKMDDEV-W-SKNNIINSVCESEPEKQOIVIRKEVS 528
 QY 515 CCEPCVPCGAGTFLNKSGLYRCOPCGKEWAPBEGSOTCPPRVYVFLALREHTSWVLLAAN 574
 DB 529 CWTCTPCCKENEYV--FDEYTCACOLGSMPTDGLCGLIPQYLRMDPDEPIAAVFA 586
 QY 575 TLLTLLLTAGLFAMHLLTPVRSAGRLCFMLGSLAAGSSLYGFGEPAPRACLIR 634
 DB 587 CIGLALTLVYVYFIITRPTPVAKSSRELCTITLIGLGLYCTFCLIAKPRQIYCYLO 646
 QY 635 QALFALGFTIFLSCUTVRSFOLIIF-----KFTVKPTFYHAWYONHAGLFWMISSAA 689
 DB 647 RIGIGSPMSYSALVTCKNRIARLAGSKKICTKPRFMSACQDLVLAFLITICL----- 702
 QY 690 QLLICLTWLVWTPLEPARYQRPPLH--VMLECTETNSLGFILAFYNGLLSISAFASY 747
 DB 703 QGIIYALFIMEPPDMDHY-----PSIREYLLICNTFN-LGVVTPLENGLLISCTFYAF 758
 QY 748 LKQDLPENNEKCVYFSLFNFSWIAFTTASVDKYLRAANMAALS-SLSGPEG 806
 DB 759 KTRNVPANNEKAYIAFWYTCIIMLAF--VPIYGSNRYKITIMCESVLSATVALGC 815
 QY 807 YFLPKCYVILCRPDNSTEHPQAS 830
 DB 816 MEVPKYYIILAKPERVRSATFTS 839

RESULT 8
 MGR1_HUMAN STANDARD: PRT: 1194 AA.
 ID MGR1_HUMAN Q13255: Q13255: Q14757: Q14758:
 AC Q13255: Q13255: Q14757: Q14758:
 DT 01-NOV-1997 (Rel. 35, Created)
 DF 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 1 precursor.
 GN GRM1 OR GPRC1A OR MGLUR1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96029774; PubMed=7476890;
 RA Desai M.A., Burnett J.P., Mayne N.G., Schoepp D.D.;
 RT "Cloning and expression of a human metabotropic glutamate receptor 1

RT alpha: enhanced coupling on co-transfection with a glutamate transporter.";
 RL Mol. Pharmacol. 48:648-657(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97231349; PubMed=9076744;
 RA Stephan D., Bon C., Holzwarth J.A., Galvan M., Pruss R.M.;
 RT "Human metabotropic glutamate receptor 1: mRNA distribution, chromosome localization and functional expression of two splice variants".
 RL Neuropharmacology 35:1649-1660(1996).
 CC -I- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM. MAY PARTICIPATE IN THE CENTRAL ACTION OF GLUTAMATE IN THE CNS, SUCH AS LONG-TERM POTENTIATION IN THE HIPPOCAMPUS AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS. STRONGEST, TO MGLUR5.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U31215; AAA87843.1; -;
 DR EMBL: U31216; AAA87844.1; -;
 DR EMBL: L76627; AAB05337.1; -;
 DR EMBL: L76631; AAB05338.1; -;
 DR GCRDB: GCR_1825; -;
 DR GCRDB: GCR_1826; -;
 DR GCRDB: GCR_1982; -;
 DR GCRDB: GCR_1983; -;
 DR MIM: 604473; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR00337; GPCR_Mgr.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_REC_P3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_REC_P3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_REC_P3_3; 1.
 DR PROSITE: PS00981; G_PROTEIN_REC_P3_4; 1.
 DR PROSITE: PS00981; G_PROTEIN_REC_P3_5; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Multigene family; Alternative splicing.
 FT SIGNAL 1 18
 FT CHAIN 19 1194
 FT DOMAIN 19 592
 FT TRANSMEM 593 615
 FT TRANSMEM 616 629
 FT TRANSMEM 630 650
 FT TRANSMEM 651 661
 FT TRANSMEM 662 680
 FT TRANSMEM 681 706
 FT TRANSMEM 707 727
 FT TRANSMEM 728 750
 FT TRANSMEM 751 772
 FT TRANSMEM 773 785
 FT TRANSMEM 786 808
 FT TRANSMEM 809 814
 FT TRANSMEM 815 840
 FT DOMAIN 841 1194
 FT DOMAIN 1014 1035
 FT DOMAIN 1067 1081
 FT DOMAIN 1095 1130
 FT DOMAIN 1142 1194
 FT CARBOHD 98 98
 FT CARBOHD 223 223

POTENTIAL.
 METABOTROPIC GLUTAMATE RECEPTOR 1.
 I (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 II (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 III (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 IV (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 V (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 VI (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 VII (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 GLN/PRO-RICH.
 GLN/PRO-RICH.
 ASP/GLU-RICH (ACIDIC).
 SER-RICH.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR PRINTS: PR00248; ANF_receptor; 1.
 DR PROSITE: PS00979; G_PROTEIN_RECEPTOR_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECEPTOR_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECEPTOR_F3_3; 1.
 DR PROSITE: PS00982; G_PROTEIN_RECEPTOR_F3_4; 1.
 DR G-protein coupled receptor; transmembrane; Glycoprotein; Signal;
 KW Multigene family.
 FT SIGNAL 1 32 POTENTIAL.
 FT CHAIN 33 912 METABOTROPIC GLUTAMATE RECEPTOR 4.
 FT DOMAIN 33 912 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 588 610 I (POTENTIAL).
 FT DOMAIN 611 624 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 625 645 II (POTENTIAL).
 FT DOMAIN 646 656 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 657 675 III (POTENTIAL).
 FT DOMAIN 676 699 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 700 720 IV (POTENTIAL).
 FT DOMAIN 721 750 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 751 772 V (POTENTIAL).
 FT DOMAIN 773 785 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 786 808 VI (POTENTIAL).
 FT DOMAIN 809 821 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 822 847 VII (POTENTIAL).
 FT DOMAIN 848 912 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 912 AA; 101867 MW; 4A2F36B53A2EA5A CRC64;

Query Match 15.7%; Score 704.5; DB 1; Length 912;
 Best Local Similarity 26.9%; Pred. No. 1.1e-43;
 Matches 228; Conservative 142; Mismatches 379; Indels 99; Gaps 31;

QY 34 LPGYLLAGFPLHSGCLQVHRREYVLCDRSCSFNENGYHLFQAMRLGVEINNNSTALL 93
 DB 45 IDGDTLGLGFLPVHGRSEKPCGELK-----KEGGIHRLEAMFALDRINNDPDL 96
 QY 94 PNILGOLVDCSDSANVATLRLVSLPGHTELOG-----DLHYSPVTLAV 143
 DB 97 PNILGARIDTC--SRDTALBOSLTFV-QALLEKDTGTEVRCSSGGEPIITTKERRVGV 153
 QY 144 IGPSTNRATTAALLSPFLVPMISVYASSETLSVKROYPSFLRTIPNDKYQVETMVL 203
 DB 154 IGAGSSSVISIVANILRLFKIPQISVASTADPLSDNSRHYDFSRVPSDVTQAQAMVDIV 213
 QY 204 QKREWTISLVGSSDDYGGQGVQA-LENOATGOGICIAFKIMPSAOGDEBRQCLMRH 262
 DB 214 RALKMNVSTVYASGSGVSEVFAIKSRDEGVCAIAOSVKIPREPRAG-EFKIIRRL 272
 QY 263 LAOGATVWVVFSSRQALAVFESEVLTNLGK-VWVAASEAMALSRHTGTGPGIORIMV 321
 DB 273 LETSNAAVITFANEDIRRLLEARRANOTGHEFFWMSDSW--GSKTAPVHLHEVAEG 330
 QY 322 LGVAIOKRAVGLKAFEEAVARADKEARPCCHKSGWCS-----SNOLRECOAFAMHTMPK 377
 DB 331 AVTLPRKM--SVRGDFRFFSSRLDNNR--RNIMFAEFMEDNPHCK-----LSRHALKK 381
 QY 378 -----LKAFSMSSAYN-----AYRAVYAVAGHL-----QLGASGACSR-GRV 416
 DB 382 GSHVKKCTNRERIGODSAYEGRGVQVPIVAVMAGHLAMHRLDLCGRGLCPRMPIV 441
 QY 417 YPMOLLEOIRKVFPL-LHKQVFAFNDRDPLSVNITIAMDNNGKMPFTVGSSTWSPVQ 475
 DB 442 DGTOLLEKIRVAVNSGIAKNVITNENGDAGRDIYQURLNDSAEKTVGS--WID-H 498
 QY 476 LNIETKIOWHGKDNQVPKSVSSDCLEGHORVYTGTHHCCEGVPCGAGTFLNKSDLIR 535
 DB 476 LNIETKIOWHGKDNQVPKSVSSDCLEGHORVYTGTHHCCEGVPCGAGTFLNKSDLIR 535

DB 499 LHLRIERHMPGSGOQLPRISICLPCQGERKKTYKGMPCMHCEPCTG--YQVQVDRY 556
 QY 536 COPGKEEWADEGSCOTCPRRVVFALREHNSWILLANTLLLLCTAG-----LPA 589
 DB 557 CKTCPYDMRPENRGTG--RPPIIKLEWGSFWALP-----LFLAVGIAATLFEVITFV 610
 QY 590 WHLDPPVRSAGRLCEFLMGLSAAAGSLGFGGEPRPACLRQALFALGFTFLSCL 649
 DB 611 RYNDPIYKAGRELSTYLLGIFLCYATTFMLAEPLGTCSLRIRFLGMSISTYAL 670
 QY 650 TVRSFOLIIF--KFTKVPFFYHAWQNHAGLFFVMISSAOLLICLTFLVYVTPPLA 706
 DB 671 LTKNRIYRIEGRKRSVADPRISPASQL--AIFFLIS--LQLLGICVWFVDPSSHV 726
 QY 707 REYQ-----RPHLVMLECTEINSGLFLLAFYNGLISARFCSYLGKDPENYEA 759
 DB 727 VDFODQRLDPRFARGV-LKC-DISDSLICLLGISMILMTCTVYAAKTGVPETFNEA 784
 QY 760 KCVTFSLLENFVSWIA-----FETTSVYDGYLPAANMAGIS--SLSSGFGYFLPKCYV 814
 DB 785 KPICFTMTTICVIALFIPFIFGTISQADKLYITTLTVSVLSASVSLGMLYKPYI 844
 QY 815 ILCRPDLN 822
 DB 845 ILFHEQN 852
 RESULT 11
 ID MGR2_RAT STANDARD; PRT; 872 AA.
 AC P31421;
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 2 precursor.
 GN GRM2 OR GPCRB OR MGLUR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.,
 RL Neuron 8:169-179(1992).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 CC MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
 CC MAY MEDIANE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN
 CC SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS AND PROMINENT
 CC EXPRESSION IS SEEN IN GOLGI CELLS OF THE CEREBELLUM AND SOME
 CC PARTICULAR NEURONAL CELLS IN OTHER BRAIN REGIONS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M92075; ?; NOT_ANNOTATED_CDS.
 CC PIR: JH0561; JH0561.
 CC HSSP: P06612; 1ECL.
 CC GCRDB: GCR_0361; --.
 CC InterPro: IPR001828; ANF_receptor.
 CC InterPro: IPR000337; GPCR_Mgr.
 CC Pfam: PF00003; 7tm_3; 1.

DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECPEP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECPEP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECPEP_F3_3; 1.
 DR PROSITE; PS0259; G_PROTEIN_RECPEP_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family.
 FT SIGNAL 1 18
 FT CHAIN 19 872
 FT DOMAIN 19 567
 FT TRANSMEM 568 590
 FT DOMAIN 591 604
 FT TRANSMEM 605 625
 FT DOMAIN 626 636
 FT TRANSMEM 637 655
 FT DOMAIN 656 679
 FT TRANSMEM 680 700
 FT DOMAIN 701 725
 FT TRANSMEM 726 747
 FT DOMAIN 748 760
 FT TRANSMEM 761 783
 FT DOMAIN 784 793
 FT TRANSMEM 794 819
 FT DOMAIN 820 872
 FT CARBOHYD 203 203
 FT CARBOHYD 286 286
 FT CARBOHYD 338 338
 FT CARBOHYD 402 402
 FT CARBOHYD 547 547
 FT SEQUENCE 872 AA; 95773 MW; 1E74CABD6AD4BED9 CRC64;

Query Match 15.4%; Score 692.5; DB 1; Length 872;
 Best Local Similarity 27.5%; Pred. No. 7.9e-43;
 Matches 245; Conservative 121; Mismatches 361; Indels 165; Gaps 38;

QY 33 TLPGDYLAGLPLHSGCLQVNRPEVTLCDRSCSFNEH-GYHLEQAKRLGVEEINNSTA 91
 DB 27 TLEGDLVGLGEPVHQ-----KGGP-----AECCGVNEHRIORLEAMFALDRINRDPH 77
 QY 92 LLPNTTGLQYLVDCSDAN-----VYATRLVLSLPGCHHILQGLDILHYS--PT-V 140
 DB 78 LLPGRLAHLLIDSCSKTHALEQALDFVRLSL--RGADGSHNICPDSDYATHSAPAV 136
 QY 141 LAVIPDSTNRATTAALISPLVPMISYAASETLSYKQYPSFLRTIPNDKQVETVY 200
 DB 137 TGVIGGSYSDVSIOYANLLRLFOIPOSYASISAKLSKSRDYARYVPRPFQAKAMA 196
 QY 201 LLLQKFGTWTISLVGSSDDYQGLQVQALENQATGGGICIAFKDIMPFSQVGDERRQICM 260
 DB 197 ELIRFENNTYVSTVASEDYGEGTIEAELEEARANICVA-----TSEKVRAMSRAAF 250
 QY 261 RHLQA-----GATVVVFSSQOLARVFEESVYVLTNLGKVVAVASEANLSHITGVPI 315
 DB 251 EGVVALLQKRSARVAVLTSEDAKRELLAATQRLN-ASFTWVADGWALESV----- 303
 QY 316 QIRGAVLGLVQKRAVPGLKAFEEA-YARAD-----KEAPRCPCHKSM-----C 358
 DB 304 -----VAGERAAEGAITTELASYPISDFASYFQSLDPNNRNPNFRFEMERFHC 355
 QY 359 SNNQLCRCQAMHTMKLAFSSMSAYNAYRAYVAHGHJOLIG-----CASSG 409
 DB 356 SFRQ--RDCAAHSLRAVFEQESKIMFYVN--AVYAMAHALHNNHRAALCPWTHLCDAM 410
 QY 410 ACSRGVYPMQLLBOIHHVHLLH-----KOTVAAANDRRPLSSNITIAMDN 457
 DB 411 RPNVR-----RLYK-DEFLNVKFDAPFRPADTDEVDREGRGDIAGRNIIFTYLA 461
 QY 458 GP-KWTFVLGSSSTWSPOLNINETKIQHGRD-NOVPKSVSSCCLBCHQVWTGFIHC 515
 DB 462 GSGRRIRYQKVG--YVAE-GLITLDTSFITPASPAGPLPRAKRCSECLONEVASYQGEVC 518

QY 516 CREVCPCAGTFLNKSLSLYRCPCCKEMAPEGSOT-CEPRTVFLALREHTSWL--LA 572
 DB 519 CWLICPCPYEX--RDEFTCADCGLGW-PNASITGCFELPOEY--IRMGAMAVGVT 573
 QY 573 ANTLILLILLITAGLPAHMLDPPVRSAGRCFLMGLSANGSSSLGFECEPRPACL 632
 DB 574 IACIALATLEVLGVFVRINATPVKASGRELCYLLGGVFLCYCMTFVIAKPTAVCT 633
 QY 633 LRQALFALGFTFLSLCRFSPOLILIF--KFTKVPTEFYHAMVQNHGAGLEVISSNA 689
 DB 634 LRLRLGIAFSYCSYALTLKTRIRITGARGAGQORPE-----ISRAS 678
 QY 690 QLLICL-----TWLVVTP-----LPAREYRPHLVMECT--ETNSLGFIL 730
 DB 679 QVAILCALISGOLLIVAAWLVEAGTGKETAPER-----EVTILRCNHRDASMLG-- 730
 QY 731 AFLVGLGISAFASYLGKDLPEVNEAKCTSEFLLENFVSWIAFPFTASYDCKYLP 790
 DB 731 SLAYVLLIALCTLYAFTRKCPENFNEKFTGFTMYTTCILMLFLPIFYTSSDYRQ 790
 QY 791 ANMMAGLSLSSG--FGGYFLPKCVIICRPDLNSTEHFOASIDYTRCGS 840
 DB 791 TTMCVSYSLSSVVLGCLFAPKHLIILFQPKNVVSH-----RAPTRFGS 837

RESULT 12
 MGR3_RAT
 ID MGR3_RAT STANDARD: PRT: 879 AA.
 AC P31422:
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 3 precursor.
 GN GRM3 OR GPRC3 OR MGLUR3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors."
 RL Neuron 8:169-179(1992).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
 EXPRESSION IS SEEN IN THE NEURONAL CELLS OF THE CEREBRAL CORTEX,
 DENTATE GYRUS, AND GLIAL CELLS THROUGHOUT BRAIN REGIONS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONEST. TO MGLUR2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M92076; -; NOT_ANNOTATED_CDS.
 CC PIR: JH0562; JH0562.
 CC GCRDB: GCR_0362; -;
 CC InterPro: IPR001828; ANF_receptor.
 CC InterPro: IPR000337; GPCR_Mgr.
 CC Pfam: PF00003; 7tm_3; 1.
 CC Pfam: PF01094; ANF_receptor; 1.
 CC PRINTS: PR00248; GPCRMR.
 CC PROSITE: PS00979; G_PROTEIN_RECPEP_F3_1; 1.
 CC PROSITE: PS00980; G_PROTEIN_RECPEP_F3_2; 1.

DR PROSITE; PS00981; G-PROTEIN_RECEP_F3_3; 1.
 DR PROSITE; PS50259; G-PROTEIN_RECEP_F3_4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family.
 FT SIGNAL 1 22
 FT CHAIN 23 879
 FT DOMAIN 23 576
 FT TRANSMEM 577 599
 FT DOMAIN 600 613
 FT TRANSMEM 614 634
 FT DOMAIN 635 645
 FT TRANSMEM 646 664
 FT DOMAIN 665 688
 FT TRANSMEM 689 709
 FT DOMAIN 710 734
 FT TRANSMEM 735 756
 FT DOMAIN 757 769
 FT TRANSMEM 770 792
 FT DOMAIN 793 802
 FT TRANSMEM 803 828
 FT DOMAIN 829 879
 FT CARBOHYD 209 209
 FT CARBOHYD 292 292
 FT CARBOHYD 414 414
 FT CARBOHYD 439 439
 SQ SEQUENCE 879 AA; 98959 MW; 3E5965EDD56DEED CRC64;

Query Match 15.3%; Score 688; DB 1; Length 879;
 Best Local Similarity 25.8%; Pred. No. 1,7e-42;
 Matches 227; Conservative 139; Mismatches 355; Indels 160; Gaps 35;

QY 31 DFTLPDGYLLAGLPPLSGCLQVHRPEVTLCDRSCSFNE-HGYHLQAMLGVEETINS 89
 DB 32 EIKIESGLVGLGFPINE-----KGTGTECGRIINEDRGIGRIEAMFAIDELINKD 82
 QY 90 TALLPNTIGYQLYDYDCSDANVATRLVSL-----PGOHIELOGDL 133
 DB 83 NYLLPGVKLGVLHLDTC--SRDFTVALQSLSEFVASILTKVDEALWCPDGYALIOENPL 140
 QY 134 LHSPTPLVATIGDPSTRAATTAALSPFLVPMSTYAASTSLVKKRQSFLETPINDK 193
 DB 141 L-----TAGVIGGYSYSVSIQVANLRLFOIPQISYASTSKLSDKSYDFARTVPDFE 195
 QY 194 YQVETWVLLLOKFGMTWISIVGSSDDYGOLGVOALENOAGGICIAFKDIMPSSAQVD 253
 DB 196 YQANAMAEILRFEMWTVSVVASEGDYGETEAFEDQARLNICIAFAE-----KVG 249
 QY 254 ERMQ-----CLMRHLAQ-AGATVVVVFSSROLARVFESVLTNLTKVWVASEAMALSRL 308
 DB 250 SNIRKSYDSVIRELLOKRNARVVLFRSDSRELLIAANRVN-ASFTWASDDMGQDES 308
 QY 309 ITGVPGIQRIGVNLGVAIQRAYDGLKAFEBAYARADKEAPRCHKGSW-----C 358
 DB 309 I--VKGESEH--AVGAILTEIASHPVROFDRYFQSLN--PYNHRRMPWFERDEWQFQC 361
 QY 359 S-----SNOLCRQCAFMAHTMPKLAFFMSSSAYNAVAYAAHGLHQ-----LLCAS 408
 DB 362 SLONKRNHRROVCDHILADISNTEQ-----ESKIMFYVNAVYAAHAKHKKORTLCPTNT 416
 QY 409 GACSRGVVPMQL--EQIKVHFLHKKDVAENDR-----DPLSSYNIIAW 454
 DB 417 KLDAMKIIDLGKKLYKLYLLKIN-----TAPFNPNGASISYKFDFFGDMGKYNFNL 471
 QY 455 DMNGPKRTFVVLGSSSTWSPVQNLINETKIQWGHGKNQVPSVSSDCLGHRVVTGFHH 514
 DB 472 QQNGGKSYLKVGH--WAET--LSLDVDSIHW--SRNSVPTSCSDPCAPNEMKMGQGDV 526
 QY 515 CCEFCVCGAGTFLNKSDDLRCPCGKEWAPESQTCF--PRIVVLAALREHISW----- 568
 DB 527 CCHICICEPEYEYL--VDETCCHDCGPGWPTADLSGCIULPEDYI-----KWEDAMVIGP 580
 QY 569 VLLAANTLILLLLIGTAGLFAMHLDPVNVASAGRLCFLMLGSLAAGSGSLYGFGEPT 628

DB 581 VTICAGLFLTCIYIF--VFIKHNHTPLVKAAGRELCTYLLFGVSLSCYCMFFFIARSP 638
 QY 629 PACILROALFALGFTIFSLCTVRSFOLIIF---KSTKVPTRYHAMVQHGAGLPMVI 685
 DB 639 VICARLRIGLICTSPALCYSAILLTKTNCIARIFEDVKNCAQRPK-----I 683
 QY 686 SSAOLICL-----TWLVWTP-----LPAREYRFPHYLMECTETNSLGF 728
 DB 684 SPSSQVFLCLDLIVQVYVSWMLLETPGTRRTLPBKR-----EYILCKNNKDS--SM 737
 QY 729 ILAFLYNGLSISAFASYLCKDLPENYNKACVTSLELFNFSWIAFTTASYDGYL 788
 DB 738 LISLVDVAVILCTVAVAFKTRKCPENNEAKFIFGTWYTCITLWAFLEPIFYVTSDDYR 797
 QY 789 PAANMAGLSLSSGF---GGYFLPKCVILCRDPLNTEH 826
 DB 798 VQTTMC-IVSLSGFVVLGLCLFAFKVHIVLFQOKNVVTH 837

RESULT 13
 MGR3_HUMAN STANDARD; PRT; 877 AA.
 AC Q14832;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 3 precursor.
 GN GRM3 OR GPRC3 OR MGLUR3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96437205; PubMed=8840013;
 RA Makoff A., Volpe F., Leichuk R., Harrington K., Emson P.;
 RT "Molecular characterization and localization of human metabotropic
 glutamate receptor type 3".
 RL Brain Res. Mol. Brain Res. 40:55-63(1996).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X77748; CA54796.1; -.
 DR GCRD; GCR_2070; -.
 DR MIM; 601115; -.
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm.3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G-PROTEIN_RECEP_F3_1; 1.
 DR PROSITE; PS00980; G-PROTEIN_RECEP_F3_2; 1.
 DR PROSITE; PS00981; G-PROTEIN_RECEP_F3_3; 1.
 DR PROSITE; PS50259; G-PROTEIN_RECEP_F3_4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family.
 FT SIGNAL 1 20
 FT CHAIN 21 877
 FT DOMAIN 21 574
 METABOTROPIC GLUTAMATE RECEPTOR 3.
 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 575 597 I (POTENTIAL).
 FT DOMAIN 598 611 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 612 632 II (POTENTIAL).
 FT DOMAIN 633 643 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 644 662 III (POTENTIAL).
 FT DOMAIN 663 686 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 687 707 IV (POTENTIAL).
 FT DOMAIN 708 732 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 733 754 V (POTENTIAL).
 FT DOMAIN 755 767 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 768 790 VI (POTENTIAL).
 FT DOMAIN 791 800 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 801 826 VII (POTENTIAL).
 FT DOMAIN 827 877 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 877 AA; 98619 MW; 66F28663CE35F740 CRC64;

Query Match 15.2%; Score 684; DB 1; Length 877;
 Best Local Similarity 25.7%; Pred. No. 3,3e-42;

Matches 225; Conservative 140; Mismatches 361; Indels 150; Gaps 33;

DB 30 EIKIGDGLVGLGLPINE-----KGTGEEGRINEDGRIQRLFAIDKIND 80
 QY 31 DFLPDYVILAGLPPLHSGCLQVRRHREVTLLCDRSCFNE-HGYHLFOAMRLGVEEINNS 89
 DB 90 TALLNITNGYLDYVCSNANVATRLSL-----PGCHHIELQGL 133
 DB 81 DYLLEGVKLVGHILPTC--SRDTYALDLSLEFVRASLTKVDAEYKCPDGSYAIQENIPL 138
 QY 134 LHYSPFLVAVIPDSTNRATTAALLSPFLVPMISYASSETLSYKQPSFLRTIPNDK 193
 DB 139 L-----IAGVIGSYSSVSIQVANNLRLEQIQAISASLSKSRDYARATPPPF 193
 QY 194 YQVETMVLKLFQGTWISLVSSDDYGLQVDALENOATGGCICAFKDIMFSAQVD 253
 DB 194 YQAKMAELREFNMTYVSTVASEGYGETGIEAEQEARLNICIATAE-----KVGK 247
 QY 254 EMO-----CLMHLAO-AGATVYVVFSSQOLARVFESVVLNLTKYVAVASEAMLSH 308
 DB 248 SNIRSYSDYIELLOKPNARVAVLEPMSDSRELIATAASRAN-ASFTVAVDGMGAQDS 306
 QY 309 ITGVPGIORIGVAVLOKRAVPGIKAFEEAYARADKEAPPCPKGSV-----C 358
 DB 307 I--IGSEHV--AYGAILTELASQPVRODRYFQSLN--PYNHNNPWFDFWQKFCQ 359
 QY 359 S---SNQCRBEOAMAHMFKLAFSSMAYN-----ATRAYAVAHGLHQ---LL 404
 DB 360 SLQNRNHRVCDKHLAID-----SSNYEOKIMFVVNAVAYAMAHALHKMORTJC 410
 QY 405 GCASGACSGRGRVYPMOLLEOIKHVFHLKDTVAFENDNDP-----LSYNI 452
 DB 411 PWTYLCIDAMKILDG---KKLYKDYLIKNTAFAPNPKNADSIYKFTFGGOMGYNYF 467
 QY 453 AWDWNGPKMTFTVLGSSTWSVPOVLINETKIQMHGKNDQPKSVSSDCLGSHORVVTGF 512
 DB 468 NFQVNGKSYLYKVGH--WAET-LSLDVNSIHW--SRNSVPTSCQSDCAPNEMKMMQGG 522
 QY 513 HNCCEVPCGAGTFLNKSNDLYRCOPCKEENAPBESQTCFRTYVFLAREHTSM---- 568
 DB 533 DVCWCWICIPCPEYEYL--ADEFTCMCGSGOMPTADLTCTCYDLPEDY--IRWEDAWAIGP 578
 QY 569 VLLAANTLILLLGLAGFAMHLDTPVYRAGRGICFLMGLSLAAGSSLYGFEEPR 628
 DB 579 VTIACIGFCTCMVYF--VFIAHNTPLVKASGRELCYLLGVGISTCMTFFFAKPP 636
 QY 629 PACLRQALFALGFTIFSLCTVRSFOLITIF--KFSIKVTFYFAYVQNHGAGLFVMI 685
 DB 637 VICALRRIGLSSFAICYSALLTKINCIAKITFDGKNGKNGAKRKF-----I 681

QY 686 SSAADLLICLTWLV-----VWTPLPAREYORFP-----HVMLECEFTNSLGFIAPL 733
 DB 682 ISSGLSITGLGLIVQVWVWLLLEAPGTRRYTLAEKRETVILKCNKDS-SMISL 740
 QY 734 YNGLSISAFASYLCKDLPENYNEAKCYTFSLLNFVSWIAFFTTASYDCKYLPAAAM 793
 DB 741 YDVILVILCTVYAFATRKCPENNEAKFTGFTWYTCIIMAFLEPIFYVSSDYRVQTT 800
 QY 794 MAGLSLSSGF---GGYFLPKCYVILCRPDLNSTEH 826
 DB 801 MC-IVSLSGFVVLGLFAPKVIHILLFPQKNVYH 835

RESULT 14
 MGR2_HUMAN
 ID MGR2_HUMAN STANDARD; PRT; 872 AA.
 AC 014416;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 2 precursor.
 GN GRM2 OR GPRC1B OR MGLUR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RX MEDLINE=95346007; PubMed=7620613;
 RA Flor P.J., Lindauer K., Putner I., Ruegg D., Lukić S., Knöpfel T.,
 RA Kuhn R.;
 RT "Molecular cloning, functional expression and pharmacological
 RT characterization of the human metabotropic glutamate receptor type
 RT 2,"
 RL Eur. J. Neurosci. 7:622-629(1995).
 CC -I- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 CC MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
 CC MAY MEDIATE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN
 CC SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- TISSUE SPECIFICITY: WIDELY EXPRESSED IN DIFFERENT REGIONS OF THE
 CC ADULT BRAIN AS WELL AS IN FETAL BRAIN.
 CC -I- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR3.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L35318; AAA76855.1;
 CC GCRDB: GCR 1846;
 CC MIM: 604093;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMBR.
 DR PROSITE: PS00979; G_PROTEIN_RECPE_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECPE_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECPE_F3_3; 1.
 DR PROSITE: PS0259; G_PROTEIN_RECPE_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; glycoprotein; signal;
 KM Multigene family.
 FT SIGNAL 1 18
 FT CHAIN 19 872 METABOTROPIC GLUTAMATE RECEPTOR 2.
 FT DOMAIN 19 567 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 568 590 I (POTENTIAL).
 FT DOMAIN 591 604 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 605 625 II (POTENTIAL).

```

FT DOMAIN 626 636 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 637 655 III (POTENTIAL).
FT DOMAIN 656 679 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 679 700 IV (POTENTIAL).
FT DOMAIN 701 725 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 726 747 V (POTENTIAL).
FT DOMAIN 748 760 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 761 783 VI (POTENTIAL).
FT DOMAIN 784 793 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 794 819 VII (POTENTIAL).
FT DOMAIN 820 872 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 872 AA; 95507 MW; 058608C3C701E9D CRC64;

```

Query Match 14.8%; Score 666.5; DB 1; Length 872;
 Best Local Similarity 27.4%; Pred. No. 6.3e-41;
 Matches 242; Conservative 126; Mismatches 366; Indels 147; Gaps 41;

```

QY 33 TLPGDYLAGLPLHSGLQVRRHREVTLCDRSGS-FNEH-GYHLFQAMRLGVEEINNST 90
D 27 TLEGDVLVGLGFVPHQ-----KGGPA-----EDCGPVNEHGIORLEAMLFALDRINRDP 76
QY 91 ALPNITLGYLYVCSDSAN-----VYATLRVLSLPCOHITELGDLIHT--SPT- 139
D 77 HLLGVRRGAHILDCSCTDRHALDQALDFVRASLS-RCGADSRSRHCPCPGSYATHDADPTA 135
QY 140 VLAVIGPSTRRAATTAALLSPFVPMISVASSSETLSVKROYPSFLTPTINDKYQVETM 199
D 136 ITGYIGGSYSVSQVAMNLFRLFOIPQISTASTSKLSDKSRIDYFANTVPDPFOQAKAM 195
QY 200 VLLQKGTWISLVGSSDDYGOLGVOALENQATGCGICAFKDIIMPESAOVDERMOC 259
D 196 AEILRFNMWTVSTEASGDYGERGIEAFELARARNICVA-----TSEKVGAMSRMAA 249
QY 260 MRHLAQA-----GATVVVYSSRQLARVEFVSVLTNLTKGVVWASEM-ALSHITTVP 313
D 250 FEGVRRALLOKPSARVAVLFTSRDARLAAOSRLN-ASETWASDGMGLAESVVAAGE 308
QY 314 GIORIGWLVGAIOKRAVPGKAFEEAVARAD--KEAPRPHKGSW-----CSSNOLCRE 366
D 309 GAAL-----GATITELASP-ISDRASYFQSLDPMNNSNPFREFWEDRFCSFRQ--RD 361
QY 367 QQAFMAHTMPLKAFSSSSAYNAVAVAHGLHOLGSCAGSCSRGVYPMOLLEQIH 426
D 362 CAASHLRAVPPEQESKIMFYVN---AVYAMAHALHNMHRALCPMTTR-----LCDAMR 411
QY 427 KVH-FLHKD---TVARNDRDPLSSYITIAMDNKPK-----WTFVLLGSSST----- 470
D 412 PVNRRRLKYKDVLVNKKDAPPRPADTNEVFRDREGDIGRNYFTYLRASGRARYQKV 471
QY 471 --NSPVOLINRETKIOMHKD--NCPVKSVCSSDCLGHRVAVTGFHHCFCFVPCAGT 527
D 472 GYMAF-GITLDTSLIPASPSAGPLASRCSEPCLOKVKVQREYVCWMLCICQPPLEY 530
QY 528 LNKSDLYRCOPCKEEMAPESQPT-CEPRVVFALAREHTSWLT--LAANTLLLLLLGT 584
D 531 --RLDEFACDCCGLGYM-PNASTLGCFELPOEY--IRMGDAWAVGPVTIACIGALATL 585
QY 585 AGLRAMHLDPVVSAGRCCELMGLSAASSGLYFGFPTPACILRQALFALGFTI 644
D 586 LGVTVRHATPVKASKELCYILLGVFLCYMTFTIAPSTAVCTLRRLGLGTAFSV 645
QY 645 FLSCITVRSFOLITIF--KFTKVPTFEYHAMVONHGAGLVMISSAQLDICTL----- 695
D 646 CYSALLTRTNRIARIFGARAGACRPR-----ISPAQVACIALISGQ 690
QY 696 -----TWLVVTP-----LPAREYQRPPIVMECT--ETNSLGLIATLYNGLISISA 742

```

```

D 691 ILIVVAMLVPEAPGTGKETAPER-----EVLTRCNHRDASMLG---SLAYVLLIALC 742
QY 743 FACSVLGDLPEAPYKACVPSFLNFVSWIA-----EFTTASYDGKILPANNMAGIS 798
D 743 TLVANTRKCPENENAFIGFTMTTCITWALLPIFYVSSDY---RVQTTMCVSVS 799
QY 799 -SLSSFGGYFLPKCYVILCRPDLNSTEHFQASIDYTRRGS 840
D 800 LSGSVVLGCLFAPRKHLIIFQPKNVSH-----RAPTSREGS 837

RESULT 15
MGR6_HUMAN
ID MGR6_HUMAN STANDARD; PRT; 877 AA.
AC O15303;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 6 precursor.
GN GRM6 OR GPRC1F OR MGLUR6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97358610; PubMed=9215706;
RA Hashimoto T., Inazawa J., Okamoto N., Tagawa Y., Bessho Y., Honda Y.,
RA Nakanishi S.;
RT "The whole nucleotide sequence and chromosomal localization of the
RL Eur. J. Neurosci. 9:1226-1235(1997).
CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR4.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; U82083; AAB82068.1; -.
DR GCRDB; GCR_2607; -.
DR MIM; 604096; -.
DR InterPro; IPR001828; ANP_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
DR PROSITE; PS0259; G_PROTEIN_RECP_F3_4; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multi-gene family; Vision.
FT SIGNAL 1 24
FT CHAIN 25 877
FT METABOTROPIC GLUTAMATE RECEPTOR 6.
FT DOMAIN 25 585
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 586 608
FT I (POTENTIAL).
FT DOMAIN 609 622
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 623 643
FT II (POTENTIAL).
FT DOMAIN 644 654
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 655 673
FT III (POTENTIAL).
FT DOMAIN 674 697
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 698 718
FT IV (POTENTIAL).
FT DOMAIN 719 748
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 749 770

```

```
FT DOMAIN 771 783 CYTOPLASMIC (POTENTIAL).
FT TRAMSMEM 784 806 VI (POTENTIAL).
FT DOMAIN 807 819 EXTRACELLULAR (POTENTIAL).
FT TRAMSMEM 820 845 VII (POTENTIAL).
FT DOMAIN 846 877 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 877 AA: 95436 MW: AECDF91EBDA5994F CRC64;
```

Query Match 14.8%; Score 666.5; DB 1; Length 877;
Best Local Similarity 25.7%; Pred. No. 6.3e-41;
Matches 232; Conservative 139; Mismatches 377; Indels 155; Gaps 33;

```
QY 11 LQLLISCCWAFACHSTRESSPDLFLPEDYLLAGLFLPLSGCLQVRRHREVTLCDRSCS--F 68
DB 12 LVALLPLAWLAQGLARAGSVRLAGGLTLGLFLPVHAR-----GAGGRACGPLK 61
QY 69 NEHGVLFPQAMRLGYEINNSTALLPNTILGYQLYDVCSDSANVATLRYLSLPGQHIE 128
DB 62 KEGGVRLAEAMLYALDRVNADELLPGVRLGARLLDTC--SRDTYALDAQALSFV-QALIR 118
QY 129 LOGD-----LHYSF--TVLAVIGPDSNRRAATTAALSPFLVPMISTYASS 173
DB 119 GRGDDGDEVGRCPGVPPLRPAPERVAVVSASSVSIMAVNLRLFAIPQISYASTA 178
QY 174 ETLISVKRQPSFLRTIPNDKYQVETVLLQKRGWISLVGSSDDYGOLGVOA-LENQA 232
DB 179 PELSDSTRIDFSRVVPSPSYQAAWVDIVRALGMNTVSTLASSENGESGEAFVQISR 238
QY 233 TGOGICIAFKDIMPFSAQYDERMOCIMRHLAAGATVVVVSRRQARVFESVVLTNL 292
DB 239 EAGVCVIAQSIKIPREPKEG-EFSKVIRRLMETPNRAGIILFANEDDIRVLEAARQANL 297
QY 293 TGR-VWVASEANALSHHTGVEIQIRIGVYLAIOKRAVPGIKAFEEAVARADKAPRP 351
DB 298 TGHFLMVGSDSWGAK--TSPILSLDEVAAGAITILPKRA--SIDGFQYFMRSLNNR- 352
QY 352 CHKGSW-----CSSNQLCREQAQAFMAHTMPRLKAFMSMSAYN----- 388
DB 353 --RNIWFAEFWEENENCKLITSSGTOSSDSTRKCTG-----EERIGNDSTIEDEGKVO 402
QY 389 -AYRAVYVAHGLHQ-----LLGCASGAC-----SRGRVYPMQLLEQIHKVHFLHKDT-V 437
DB 403 FVIDAYATIAHLHSMHQALCPGHTGLCPAMEPTDGRM---LLOYTIRAVRENGSAGTPY 458
QY 438 AFENDRDLPLSINIITAMD-WNGPKMTFTVLGSGTSPVOLININETKIOMHGKNDQVPSKY 496
DB 459 MENENDADAGRVDIFQYQATNGSASSGQYQAVQOMAEF-LRLDVEALOWSGDPHEVPSSL 517
QY 497 GSSDCLGEGHORVYTGHHCCFECVPCGAGTFLNKSDLYRCOPGCKEEMAPESQOTCEPRT 556
DB 518 CSLPCGPGERKKMKVGPCCMCCEACDGYRF-QVDEFTCEACPGDMRPTPNHTGCRPTP 575
QY 557 VYFLALREHTSWLLAANTLLILLILG-----TAGLFAMHLDTPVYRSAGRLCEFLMG 610
DB 576 VYRL-----SWSSPWAAPRLLAVIGIATTVATFVRYNNTPIYRASGRSLSYLLLT 629
QY 611 SLAAGSGSLYGFGEPTPACILROALFALGFTIPLSCLTVRSFOLIIF--KSTKVP 667
DB 630 GIFLIYATIFELWAEPAVCAARLFLGITLISYALLTKNRIYRIFEQGRSVTPP 689
QY 668 TFYHAMVQNHGAGLFVMISSAQLLIC-----LFWLVVMTPLPAREYQRPFLY---- 716
DB 690 PF-----ISPTSQLVTFESLTSIQVVGMIAM-LGARP-----PHSVIDYE 728
QY 717 -----MLCCTEINSIGFIILAFYNGILSISAFACSYLGKDLPENYNEACVTF 764
DB 729 EORTVDEPARGVLKC-DMSDSLISGLCTGSLILAMTCTYATKARGVPETFEAKPIGF 787
QY 765 SLLENFVSWIA-----FTTASYVDGKYLPAANMAGLS-SLSSGFGGYFLPKCYVILCRP 819
```

```
DB 788 TWTTCIIMLAFVPIFFGTAAQSAEKIYIQTTLTVSLASASVSLGMLYVPKTYVILFHP 847
QY 820 DLN 822
DB 848 EQN 850
```

Search completed: June 11, 2002, 10:26:28
Job time: 471 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 09:32:26 ; Search time 43.75 Seconds
(without alignments)
1847.112 Million cell updates/sec

Title: US-09-819-946-2

Perfect score: 4493

Sequence: 1 MLCIARLVGLQLISCCWA.....NSTEHFQASIDYTRCGST 841

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1147	25.5	858	2 JC7683	taste receptor T1R
2	1142	25.4	1078	2 A56715	calcium receptor (
3	1130	25.2	1088	2 B56715	calcium receptor (
4	1126.5	25.1	1085	2 S40476	Ca(2+)-sensing rec
5	1123	25.0	1079	2 I59362	calcium/polyvalent
6	741.5	16.5	1180	2 JC2132	metabotropic gluta
7	741.5	16.5	1212	2 JC2131	metabotropic gluta
8	727	16.2	1171	2 A42916	metabotropic gluta
9	727	16.2	1199	2 A41939	G protein-coupled
10	708.5	15.8	912	2 JH0563	metabotropic gluta
11	703	15.6	1218	2 S71376	glutamate receptor
12	692.5	15.4	872	2 JH0561	metabotropic gluta
13	688	15.3	879	2 JH0562	metabotropic gluta
14	688	15.3	879	2 JC7160	metabotropic gluta
15	664	14.8	871	2 A46742	metabotropic gluta
16	652	14.5	908	2 I49142	metabotropic gluta
17	619.5	13.8	915	2 A49874	metabotropic gluta
18	590.5	13.1	999	2 T27628	hypothetical prote
19	517.5	11.5	1267	2 T21340	hypothetical prote
20	398	8.9	551	2 T30806	metabotropic gluta
21	263	5.9	1099	2 T16283	hypothetical prote
22	254	5.7	923	2 F84732	probable ligand-ga
23	253	5.6	921	2 T51136	ionotropic glutama
24	218	4.9	1039	2 T45779	probable glutamate
25	192.5	4.3	933	2 C96495	probable ligand-ga
26	192	4.3	925	2 T06128	hypothetical prote
27	182	4.1	953	2 E84732	probable ligand-ga
28	181.5	4.0	958	2 T02741	probable ligand-ga
29	178.5	4.0	925	2 T51133	ligand gated chann

30	178.5	4.0	951	2 T51132	probable glutamate
31	178.5	4.0	975	2 A84550	probable ligand-ga
32	178	4.0	912	2 T51131	ligand gated chann
33	178	4.0	997	2 S33754	glutamate receptor
34	176	3.9	960	2 JE0356	gamma-aminobutylic
35	169.5	3.8	950	2 T51134	ionotropic glutama
36	166	3.7	962	2 D86186	hypothetical prote
37	165	3.7	965	2 T51244	N-methyl-D-asparta
38	163	3.6	938	2 A46612	N-methyl-D-asparta
39	163	3.6	976	2 T51137	ionotropic glutama
40	160	3.6	941	2 T51135	ligand-gated chann
41	158	3.5	885	2 JN0339	N-methyl-D-asparta
42	158	3.5	901	2 JN0337	N-methyl-D-asparta
43	158	3.5	922	2 JN0338	N-methyl-D-asparta
44	158	3.5	938	2 S19710	N-methyl-D-asparta
45	156	3.5	938	2 S21104	N-methyl-D-asparta

ALIGNMENTS

RESULT	1	Query Match	25.5%	Score 1147;	DB 2;	Length 858;
JC7683	1	Best local similarity 32.6%; Pred. No. 2e-79;				
Sequence: Mus musculus (house mouse)		Matches 271; Conservative 150; Mismatches 356; Indels 54; Gaps 16;				
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001						
C:Accession: JC7683						
R:Kitagawa, M.; Kusakabe, Y.; Miura, H.; Ninomiya, Y.; Hirao, A.						
Biochem. Biophys. Res. Commun. 283, 236-242, 2001						
A:Title: Molecular genetic identification of a candidate receptor gene for sweet tast						
A:Reference number: JC7683; MUID:21222875; PMID:11322794						
A:Contents: tongue						
A:Accession: JC7683						
A:Molecule type: mRNA						
A:Residues: 1-858 <KIT>						
A:Cross-references: DDBJ:AB049994						
C:Comment: This protein, a seven-transmembrane receptor, serves as a mediator of the						
A:Gene: tlr3						
A:Map position: 4						
C:Keywords: transmembrane protein						
Query Match						
Best local similarity 32.6%; Pred. No. 2e-79;						
Matches 271; Conservative 150; Mismatches 356; Indels 54; Gaps 16;						
QY 29 SPDFTLPGDYLLAGLFPPLHS--GCLQVRHREPTVLCDSGSEFNGHYLFOAMRLGYEE 85						
DB 26 SQQFAQGDYILGGLFPLGSTFEATLNRQPNISPCNR--FSPGLGLFAMAMKMAVEE 82						
QY 86 INNSTALLPNTLLGQYLDVCSDS-ANYATLRVLSLGGQHIELQGLLHSPVYAVI 144						
DB 83 INNGSALLPGLRLGYDLDTCSSEPVYTKSSLMFLAKYSOSIAYCVYTOQPPVLYI 142						
QY 145 GPDSTNRATTAALSPVLPVPISYAASSETLSVRYRPSFLRTIPNRYQVETWLLQ 204						
DB 143 GPHSELALITGKFFSFLMPQVYSASASMDRLSDRETPSPFRITVPSDRVQLQAVYTLQ 202						
QY 205 KFGWTWISLVSGSDDYGLQVQVALENQATGQICIAFDIMFSAQVDE--RMQCLMRH 262						
DB 203 NFSWMVVAALGSDDDYGGREGLSIFSSLANANGICIAHGVLVQHDTSQQLKVDVLRQ 262						
QY 263 LAQAGATVYVYSSRQLARVPFESVVLNLNCKVAVASANAHLNHTIGVQICIQVGL 322						
DB 263 VNQSVQVYVYFASARAVYSLEYSYIHGLSPKVVAVASESWLTSVLTLPINAVGVYL 322						
QY 323 GVAIQKRAVPGKAFEEVYARADKAPRCHKSGMSS-----NOLCREQAF 370						
DB 323 GFLQGALIPERSHYETHAL---AADP-----ACFASLMELEENHMGRCRCDDI 375						
QY 371 MAHTPKLKAFSSMSA-----YNNAYRAVYAVAHGLHQLLGASGAGCS-GRGVYPMOLLE 423						

Db 376 MTONSSGLQNLISAGLHQIFATYAAVYSAQAALHNTLQCNVSHCHVSEHYLPQMLE 435
 Qy 424 QIHKVFLHKPTVAFNDRDPLSSYNIITAMPNGPKFTYVSGSTWSPVOLINETKI 483
 Db 436 NKYNSFHARDLLOFDAGNDMEYDLKMWQSPPTVLHYVTNGT---LQLOQSKM 492
 Qy 484 QNHKDNQPKSVCSDDLGHQVVTGFHHCCEFCVPCAGTFLNKSDLYRCQPCKEE 543
 Db 493 YWPG--NQVPVSCSRQCKDGQRRKGFHSCCYDCVCKAGSYRKHPDDFTCTPCNQDQ 550
 Qy 544 WAPESQTCFPRVYFLARHHTSWLLAANTLLLLLTGTLGFAMHIDTPVRSAGG- 602
 Db 551 WSPENSTACLPKPKFLANGEPVLSLLLCIVGLALALGLSVHMDSPVQASGGS 610
 Qy 603 RLCE--LIMGLSIAAGSGSYGFEGETPRACLLRQALFALGFTFLSCLTVRSFOLITF 660
 Db 611 QRCFGILGICLCL---SVLLFPGRRSSASCLAQOPMAHLPLTGCLSTFLQAAETFY-- 665
 Qy 661 KESTKVPTEYHAMVONHAGLF---VMISAAQLLICTLWLVWTPLPAREYQRPBHV 716
 Db 666 --ESELPLSMAMWLCISYLRGLMAMLVLLATFEVALCAMYLIAFPREVYDMSVLPTEY 723
 Qy 717 MECEFTNSLGLTFLATLYGLSISAFASYLGKDPENYNEAKCTFSLTFEVSIAF 776
 Db 724 LEHCYRWSVSLGLVHITAMLAFLCIGTFLVQSQPGRRNRARGLTFMLAFITWSEF 783
 Qy 777 FETASVYDCKYLPAAANMAGLSLSSGFGYFLPKCYVILCRPDLNTEHEF 827
 Db 784 VPLLAVQVAYQPAVOMGALIVCALGILVTFHLPKCYVILMPKLTQGEF 834
 RESULT 2
 A56715
 calcium receptor (clone pHPCar-4.0) - human
 C.Species: Homo sapiens (man)
 C.Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 01-Dec-2000
 C.Accession: A56715; S49341; A49419; B49419; C49419
 R.Garrett, J.E.; Capuano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert, S.C.
 J. Biol. Chem. 270, 12919-12925, 1995
 A.Title: Molecular cloning and functional expression of human parathyroid calcium recept
 A.Reference number: A56715; MUID:95279439
 A.Accession: A56715
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-1078 <GAR>
 A.Cross-references: GB:U20759; NID:9683744; PIDN:AAA86503.1; PID:9683745
 R.Pearce, S.H.S.; Thakker, R.V.
 submitted to the EMBL Data Library, August 1994
 A.Reference number: S49341
 A.Accession: S49341
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-180, 'Q', 182-989, 'R', 991-1078 <PEA>
 A.Cross-references: EMBL:X81086
 R.Pollak, M.R.; Brown, E.M.; Chou, Y.H.; Hebert, S.C.; Marx, S.J.; Steinmann, B.; Levi,
 Cell 75, 1297-1303, 1993
 A.Title: Mutations in the human Ca(2+)-sensing receptor gene cause familial hypocalciuric
 A.Reference number: A49419; MUID:94094324
 A.Accession: A49419
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 178-192 <POL>
 A.Experimental source: family N
 A.Note: sequence inconsistent with nucleotide translation
 A.Note: sequence modified after extraction from NCBI backbone
 A.Note: 186-Arg mutation is associated with familial hypocalciuric hypercalcemia and nec
 A.Note: sequence extracted from NCBI backbone (NCBIN:142453)
 A.Accession: B49419
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 289-303 <PO2>
 A.Experimental source: family E
 A.Note: sequence modified after extraction from NCBI backbone

A.Note: 298-Lys mutation is associated with familial hypocalciuric hypercalcemia and
 A.Note: sequence extracted from NCBI backbone (NCBIN:142455)
 A.Accession: C49419
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 788-802 <PO3>
 A.Experimental source: family J
 A.Note: sequence modified after extraction from NCBI backbone
 A.Note: 796-Trip mutation is associated with familial hypocalciuric hypercalcemia and
 A.Note: sequence extracted from NCBI backbone (NCBIN:142457)
 C.Keywords: glycoprotein; receptor; transmembrane protein

Query Match 25.4%; Score 1142; DB 2; Length 1078;
 Best local similarity 31.0%; Pred. No. 6.3e-79;
 Matches 276; Conservative 163; Mismatches 358; Indels 92; Gaps 22;

Qy 16 SCCW--AFACHSTESSPD--FTLPEDYLLAGLFLPHSGCL---OVRHREYVTLCDRSC 66
 Db 5 SCCWVLLATLWHTSAVGPQRAOKKGDILIGLFFHFGVAARDKDDPKSRPESVEICR-- 62
 Qy 67 SFNKGHFLFOAMRLGVEEINNSTALLPNTLGYOLDYCDSPASANYATLRLVSLPGOH 126
 Db 63 -YHFGFRLQAMIFALIEINSSPALPNTLGYRLFTGCTNVSK--ALEATLSFYAOKK 119
 Qy 127 IELOG-----DLIHSPTVLAVIPDSTNRATTAALSPFLVPMISYASSETLSVKRQ 181
 Db 120 IDSINDEFCNSENHPSITLAVGATGSGVSTAVANLLLFITPQSVYASSRLSNKQ 179
 Qy 182 YPSFLTITNDKXQVETMYLLLOKFGWTISLVGSDDYGOGLQVQALENOATGOGICIAF 241
 Db 180 FKSFLTITNDKXQVETMYLLLOKFGWTISLVGSDDYGOGLQVQALENOATGOGICIAF 239
 Qy 242 KDIIMPQAVQDEBMOCKLRHILAOAGATVYVYSSRQLARVEESVLTNLGKVVASE 301
 Db 240 SEIL--SQSDEBEIOHVEVIONSTAKIYVFPSSPDLEPLKEIVRNITGKTMLAE 297
 Qy 302 AMALSHRITGVPGIORIGVLAQRAVPGIKAF-----EEAYARADKEAPRP 351
 Db 298 AMASSSLIMPQVFNHVGSTIGFALKAGQIPGFRFLKYHPRKSVHNGFAKEFMEETPN 357
 Qy 352 CH-----KG-----SWCSSNOLCECCQAFAMHPKAKAFMSMSA 386
 Db 358 CHQEGAKGPLVDYFLRGHESGDRFSSTAEPRLCTGDEISSVEFPYDYTHLRIS 417
 Qy 387 YNAAYRAVAVAGLHOLLCC-----ASGACSR--GRVPMOQLDQIHVHFLHK--DTV 437
 Db 418 YNVIYLAIVSIAHRLDQIYCLGRGLFTNGSCADIKKVAQVILKRLRLNLTNNNGEY 477
 Qy 438 AFNDNRDPLSSYNIITAMPNGPKFTYVSGSTWSPV-----OLNINETKIOMH 486
 Db 478 TFECDGLDGNYSIIMHLSPE-----GSIVFKVGYNYAKKGERLFTNEEKILMS 531
 Qy 487 GKDQVPRKSVCSDDLCEGHQV--VYGFHHCCEFCVPCAGTFLNKSDLYRCQPCGEEMA 545
 Db 532 GSREVPFNCGRDCLAGRKGIIEBPCECEVCPGEVSDASACNKPDPDWS 591
 Qy 546 PEGSOTCFPRVYFLARHHTSWLLAANTLLLLLTGTLGFAMHIDTPVRSAGRLC 605
 Db 592 NENHTSCIAKEIEFLSWTEPFGALTLFVAVLGFILAFVLYKRNPIYKATRELS 651
 Qy 606 FLMLGSLAAGSGSLGFESEPTPRACLLRQALFALGFTFLSCLTVRSFOLITIFRFSK 665
 Db 652 YLLIFLILCCFSSLFIEIPDMTCRLQOPFGISFVICISILVKTIRVLVEE--AK 709
 Qy 666 VPTFYH--AMVONHAGLFAMISSAQLICTLWLVWTPLPAREYQRPBHVLMLEETEN 724
 Db 710 IPTSFHKKWGLNLOFLVFLCFTMQIVICVILWLTAPSPSRNOLDEDEITITFHES 769
 Qy 725 --SLGFIILAVNGLSISAFASYLGKDPENYNEAKCTFSLTFEVSIAFETT--AS 781
 Db 770 LMAIGFLIG--YTCLLAALCFPAFKSRKLPENFNNAKITTSMLLFIIVWSIFPAVMS 827

QY 782 VYDQKYLPAANMAGLSLSSGFGYFLPKCYVILCRPDLNSTEHFQAS 830
 DB 828 TY-GKFSVAEVIAMLAASFGLACIFPNKITYILLFKPSRMTIEVRCS 875

RESULT 3

calcium receptor (clone pHPCAR-5.2) - human
 C/Species: Homo sapiens (man)
 C/Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000
 C/Accession: B56715
 J:Garrett, J.E.; Capriano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert, S.C.
 R: Biol. Chem. 270, 12919-12925, 1995
 A:Title: Molecular cloning and functional expression of human parathyroid calcium receptor
 A:Reference number: A56715; MUID:95279439
 A:Accession: B56715
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1088 <GAR>
 A:Cross-references: GB:U20760; NID:9683746; PIDN:AAA8504.1; PID:9683747
 C:Keywords: glycoprotein; receptor; transmembrane protein

Query Match 25.2%; Score 1130; DB 2; Length 1088;
 Best Local Similarity 30.8%; Pred. No. 5.3e-78;
 Matches 277; Conservative 162; Mismatches 358; Indels 102; Gaps 23;

QY 16 SCCM---AFCHSTRESSPD--FTLPDYLLAGLPPLHSGCL---QVRHREVTLCRSC 66
 DB 5 SCCVLLALMTWHTSAYGPDORAKKGDILLGLPPIHGVAKADODKRESVEVCIR-- 62
 QY 67 SFNEGHYLFQAMRLGVEEINNSTALLPNTILGYOLYVCSDSANVYATLRLSPGOH 126
 DB 63 -YNRGRWLMQAMFAIEEINSSPALPRLNLTGRITDCTVSK--ALEATLSFVAONK 119
 QY 127 IELOG----DLHSPVTLAVIGPDSNTRATTAALLSPFLVPMISYASSETLSYVKRQ 181
 DB 120 IDSNLDFEFCNCSEHPISTIAVVGATGSGVSTAVANLGLFYIPQVSASSRSLSNKQ 179
 QY 182 YPSLRTIPNDKYQVETWVLLQKFGMTWISLVGSSDDYGOGLVQALENOATGOGICIAF 241
 DB 180 FKSFLRTIPNDHQATADIIETFRMNVGTIAADDDYGRPGLEKFEKEEEDIDICDF 239
 QY 242 KDIPFAQVQDERMOCIMRLHAOGATVVVVFSSROLARVFESVLTNLTKGVWASE 301
 DB 240 SELT--SQYSDERIOHVEYIONSTAKYIVFSSGPDLEPLIKEIYARNITGKIWLASE 297
 QY 302 AMALSRHTGVPGIQRIGMVGVAIQRKAVPGLKAF-----EEAVARADKEARP 351
 DB 298 AMASSSLIAMPQYPHVVGTFGIFALKAGQIPGFRFLKTVHPRKSVHNGFAKEEETFN 357
 QY 352 CH-----KG-----SWCSSNOLCREQCAQMAHTMPKLFKAFSMSA 386
 DB 358 CHLOEGAKGPPVDFTFLRGHEESGDRFSNSTARFPLCTGGENISSVETPIDYTHLIS 417
 QY 387 YNAVAYAVAHGLHQLG-----ASGACSR--GRVYPMQLLEQHKVHFLHAK--DTV 437
 DB 418 YNVTLAYSTIAHALODIYTCLPGRGLFTNGSCADIKKEAVQVHLKHLHFLTNMGQOV 477
 QY 438 AFNDNRPLSSYNTIADWNGPKMTFTVLGSSVSPV-----QLNINEFKIOWH 486
 DB 478 TFDECGDLVGYSTIINMHLSPED-----GSIVFEKGYVYVYAKKGRFLINKEKIMS 531
 QY 487 GKDN-----QVRKSVSSDCLBQHQR--VVTGFHHCFCFECVPCGAGTFLKASDLIR 535
 DB 532 GFSREPLTFLVSVLQVPEFNSCRDCLAGTRKGIIEGEPCTCFECVECPDEGSEYSDTASA 591
 QY 536 CQPCGKEMAPESQTCERPVVFLALREHNSWVLLAANTLLLLGLTAGLAFMHIDTP 595
 DB 592 CNKCPDDPMSNENTSTAKIEFLSWTEPGIALTLPAVAGITFLAFLVGLVFKFRNTP 651
 QY 596 VRSAGRLGFLMGLSLAAGSGSLYGFGEPTRPACLLRQALFALGFTIPLFSCLTVRSFQ 655

DB 652 IYKATNRELSTLLFSLCCFSSSLFIEGEPQDTCRLQPAFGISFYLICISCLIVKTRN 711
 QY 656 LIIFKSTKPYETFYH--AWQNHGAGLFVMTSSAOLILCLTWLVMPPLAREQORPH 714
 DB 712 VLVAF--AKIPSTFHRKMGVNLQFLVFLCTMQOYICVIMLYTAPSSYROLEDE 769
 QY 715 LVMLECEETN--SLGFIPLFYNGLSTISAFACVYLGKDPENYNEAKVTFSLFNEVS 772
 DB 770 IIFITCHGSLMALGFLIG--YTCLLAACFFPAFKSRKLPENNEAKFIIFMSLIFPIV 827
 QY 773 WIAFFT--ASYVDKYLPAANMAGLSLSSGFGYFLPKCYVILCRPDLNSTEHFQAS 830
 DB 828 WISFIPAVASTY--GKFSVAEVIAMLAASFGLACIFPNKITYILLFKPSRMTIEVRCS 885

RESULT 4

Ca(2+)-sensing receptor - bovine
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
 C/Accession: S40476
 R:Brown, E.M.; Gamba, G.; Riccardi, D.; Lombardi, M.; Butters, R.; Kilfor, O.; Sun, A.
 Nature 366, 575-580, 1993
 A:Title: Cloning and characterization of an extracellular Ca(2+)-sensing receptor fro
 A:Reference number: S40476; MUID:94077182
 A:Accession: S40476
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1085 <BRO>
 A:Cross-references: GB:S67307; NID:9453108; PIDN:AAB29171.1; PID:9453109

Query Match 25.1%; Score 1126.5; DB 2; Length 1085;
 Best Local Similarity 30.7%; Pred. No. 9.8e-78;
 Matches 274; Conservative 163; Mismatches 362; Indels 93; Gaps 22;

QY 14 LISCMWAFCHST---ESSPD--FTLPDYLLAGLPPLHSGCL---QVRHREVTLCDC 63
 DB 3 LYSCMWILAFSTWCTSAVGPDRQKKGDIILGLPPIHGVAKADODKRESVEVCIR 62
 QY 64 RSCSFNEGHYLFQAMRLGVEEINNSTALLPNTILGYOLYVCSDSANVYATLRLSPG 123
 DB 63 R--YNRGRWLMQAMFAIEEINSSPALPRLNLTGRITDCTVSK--ALEATLSFVA 117
 QY 124 OHTELOG----DLHSPVTLAVIGPDSNTRATTAALLSPFLVPMISYASSETLSY 178
 DB 118 ONKIDSLNDFEFCNCSEHPISTIAVVGATGSGISTAVANLGLFYIPQVSASSRSLSN 177
 QY 179 KROYPSLRTIPNDKYQVETWVLLQKFGMTWISLVGSSDDYGOGLVQALENOATGOGIC 238
 DB 178 KNOFKSLRTIPNDHQATADIIETFRMNVGTIAADDDYGRPGLEKFEKEEEDIDICDF 237
 QY 239 IAFKDIMEFSAQVQDERMOCIMRLHAOGATVVVVFSSROLARVFESVLTNLTKGVW 298
 DB 238 IDFSELI--SQYSDERIOHVEYIONSTAKYIVFSSGPDLEPLIKEIYARNITGKIWL 295
 QY 299 ASEAMALSRHTGVPGIQRIGMVGVAIQRKAVPGLKAF-----EEAVARADKEA 348
 DB 296 ASEAMASSSLIAMPDEYFHVVGTFGIFALKAGQIPGFRFLKTVHPRKSVHNGFAKEE 355
 QY 349 PRPCH-----KGSWSSNO-----LCREGCAQMAHTMPKLFKAFSM 383
 DB 356 TFNCHLOEGAGKPLPVDTFLRGHEEGGKARLSNPFARFPLCTGGENISSVETPIMDYTHL 415
 QY 384 SSAYNAVAYAVAHGLHQLG-----ASGACSR--GRVYPMQLLEQHKVHFLHAK-- 434
 DB 416 RISTNVYLAYSTIAHALODIYTCLPGRGLFTNGSCADIKKEAVQVHLKHLHFLTNMG 475
 QY 435 DTVAFNDNRPLSSYNTIADWNGPKMTFTVLGSSVSPV-----QLNINEFKI 483
 DB 476 EOVTFDECGDLVGYSTIINMHLSPED-----GSIVFEKGYVYVYAKKGRFLINDEKI 529
 QY 484 OMHGKDNQVPRKSVSSDCLBQHQR--VVTGFHHCFCFECVPCGAGTFLKASDLIRQOPCGKE 542

Db 647 ORIGISLSPAMSYALVTKTNRRIARILAGSKKICIKRPFMSACQVIAFLICI--- 703
 QY 689 AOLLCITLWVMTPLPAREYORFPHL--VMECTETNSIGFLIAFLYNGLISAFACS 746
 Db 704 -QLGIIVAFIMERPDIIMDY---PSIREVYLICNTN-LGVYTPLGYNLLISCTFAFA 758
 QY 747 YLKDLPENYNEAKCVTFSELNFVSWIAFFTASYDGKYLPAANMAGLS-SLSGGFG 805
 Db 759 KTRNVPANFNENAKYIAFTYTCIIMLAF---VPYFGSNYKIIIMCFSVLSATVALGC 815
 QY 806 GYELPKCYVILCRPDNSTEHFQAS 830
 Db 816 CMFVPRVYIILAKPERNVSATFTTS 840

RESULT 8

metabotropic glutamate receptor mglur5 - rat

C.Species: Rattus norvegicus (Norway rat)

C.Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C.Accession: A42916

R.Abe, T.; Sugihara, H.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.

J. Biol. Chem. 267, 13361-13368, 1992

A.Title: Molecular characterization of a novel metabotropic glutamate receptor mglur5 cc

A.Reference number: A42916; MUID:92317054

A.Accession: A42916

A.Status: Preliminary

A.Molecule type: mRNA

A.Residues: 1-1171 <ABE>

A.Cross-references: GB:D10891; NID:9220813; PIDN:BA00171.1; PID:d1002186; PID:g220814

A.Note: Sequence extracted from NCBI backbone (NCBIN:107749, NCBIPI:107750)

C.Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.2%; Score 727; DB 2; Length 1171;

Best Local Similarity 27.2%; Pred. No. 3.7e-47;

Matches 235; Conservative 151; Mismatches 356; Indels 122; Gaps 33;

QY 34 LPGDYLLAGFLPHSGCLOVRRHREY-TLCDRSCSF--NEHGYNHFOAMRLGVEEINNST 90
 Db 31 MGDIIIGALFVSF-----HQPVDKVKHERKCGAVRGYQIGVREAMLHLERINSDP 83
 QY 91 ALLPNTTGLYQYLVCSDSANYYATLRYLSLPGQHHEILOGDLHYS----- 137
 Db 84 TLLPNTTGLCEIRDSQWHS-----VALEDSIEIRDSLSSEEEGLVRCVDGS 133
 QY 138 -----PTVLAVIIGPDSSTRATTAALLSPFLVPMISYVASSPFLSVKQYSELTIPN 191
 Db 134 SSFRSKRPVIGYIGSSSSVAIQVONLQLENIPOIAVSATSDLSDKTLFFYFMKVPVS 193
 QY 192 DKYQVETVLLQKFGMTWISLVGSSDDYQGLGVQALENOATGOGICIA--FKDIMFESA 249
 Db 194 DAOQARAWVDYKRNMTYSAVHTEGNGESGMEAFKMSAKEGICIAHSYKITSNAGE 253
 QY 250 QVGERMOCIMHNLQAQATVYVVSRRQLARVFESVVLNLTGK--VWVASEANALSRH 308
 Db 254 QSFDFKLLKLRSHLPK--ARVVAACEGEMTVGLMLAMRRLAGFEILLGSDGMADRYD 311
 QY 309 IGVGVGIORIGVNLGVALQKRAVPGIKAFEEVYARADKRAPRCHKGSCSSNOJOREQ 368
 Db 312 VV--DGYQK-EAVGGITTLQSS--PDVKPDDYLL--KLRPETNLRNPV-----FQ 355
 QY 369 AFMAHTMP--KLKAFSM--SSAYN-----AVYAAVAVAHGLH--QL 403
 Db 356 EFWQIRFOCRLEGFQAEQENKYNKTCNSLTLRTHYVDSKMGFVNAITYSMAGLHNMOM 415
 QY 404 LGCA--SGACSR-----GHVYPMOLLQIHKVHFL--LAKDTVAFNDNRDPLSSYNIAND 455
 Db 416 SLCPGYAGLCDAKMRPIDR-----KLDSLMKTNTGVSQDMILLFDENGDSPPRYELIMNK 471
 QY 456 WNGPKWTFVYLVGSSTVPVQVLINETKTIQWKGKNOVPRKSVSSDCLSEHQRYV--TGFPH 514

Db 472 EMGKDY--FDYINVGSNDNGELMKDDEV-W-SKKNNIIRSVCSPECKGQIVIRKEYS 528
 QY 515 CCEFCPCAGTFLNKSDDYRCOPCGKEBMAPEGSCOTCPRTVVFALREHSTWLLAAN 574
 Db 529 CCWTCITPCENENY--FDEYTKACQOLGSMPTDGLGCLIPVQYLIRMDPREPIAAVFA 586
 QY 575 TLLLLLTGAGLIFAMHLDTPVRSAGRLCFMLGSLAAGSGSLYGFEPTPACILR 634
 Db 587 CGLATLAEVTVIIFYIRDTPVAKSSRELGYIILAGICLGICFCLAKPKQIYCYLO 646
 QY 635 QALFALGFIIFSLCTLVRSFOLLIF----KFSKVPFVFHAWQNHAGLFWMISSA 689
 Db 647 RIGIGISLSPAMSYALVTKTNRRIARILAGSKKICIKRPFMSACQVIAFLICI--- 702
 QY 690 QLLICTLWVMTPLPAREYORFPHL--VMECTETNSIGFLIAFLYNGLISAFACS 747
 Db 703 -QLGIIVAFIMERPDIIMDY---PSIREVYLICNTN-LGVYTPLGYNLLISCTFAFA 758
 QY 748 LKQDLPENYNEAKCVTFSELNFVSWIAFFTASYDGKYLPAANMAGLS-SLSGGFG 806
 Db 759 KTRNVPANFNENAKYIAFTYTCIIMLAF---VPYFGSNYKIIIMCFSVLSATVALGC 815
 QY 807 YELPKCYVILCRPDNSTEHFQAS 830
 Db 816 CMFVPRVYIILAKPERNVSATFTTS 839

RESULT 9

G protein-coupled glutamate receptor - rat

C.Species: Rattus norvegicus (Norway rat)

C.Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C.Accession: A41939; S15362

R.Houamed, K.M.; Kulper, J.L.; Gilbert, T.L.; Haldeman, B.A.; O'Hara, P.J.; Mulvihill

Science 252, 1318-1321, 1991

A.Title: Cloning, expression, and gene structure of a G protein-coupled glutamate rec

A.Reference number: A41939; MUID:92022526

A.Accession: A41939

A.Status: Preliminary; not compared with conceptual translation

A.Molecule type: nucleic acid

A.Residues: 1-1199 <HOU>

A.Cross-references: GB:M61099; NID:9397806; PIDN:AAA19497.1; PID:g204460

A.Experimental source: cerebellum

A.Note: Sequence extracted from NCBI backbone (NCBIPI:60785)

R.Masu, M.; Tanabe, Y.; Tsuchida, K.; Shigemoto, R.; Nakanishi, S.

Nature 349, 760-765, 1991

A.Title: Sequence and expression of a metabotropic glutamate receptor.

A.Reference number: S15362; MUID:91156047

A.Accession: S15362

A.Status: Preliminary

A.Molecule type: mRNA

A.Residues: 1-1199 <MAS>

A.Cross-references: EMBL:X57569; NID:956646; PIDN:CAA40799.1; PID:g56647

C.Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.2%; Score 727; DB 2; Length 1199;

Best Local Similarity 26.7%; Pred. No. 3.8e-47;

Matches 230; Conservative 155; Mismatches 362; Indels 116; Gaps 32;

QY 34 LPGDYLLAGFLPHSGCLOVRRHREY-TLCDRSCSF--NEHGYNHFOAMRLGVEEINNST 90
 Db 41 MGDIIIGALFVSF-----HQPVDKVKHERKCGAVRGYQIGVREAMLHLERINSDP 83
 QY 91 ALLPNTTGLYQYLVCSDSANYYATLRYLSLPGQHHEILOGDL----- 134
 Db 94 TLLPNTTGLCEIRDSQWHS-----VALEDSIEIRDSLSSEEEGLVRCVDGS 133
 QY 135 -----PTVLAVIIGPDSSTRATTAALLSPFLVPMISYVASSPFLSVKQYSELTIPN 191
 Db 134 SSFRSKRPVIGYIGSSSSVAIQVONLQLENIPOIAVSATSDLSDKTLFFYFMKVPVS 193
 QY 192 DKYQVETVLLQKFGMTWISLVGSSDDYQGLGVQALENOATGOGICIA--FKDIMFESA 249
 Db 194 DAOQARAWVDYKRNMTYSAVHTEGNGESGMEAFKMSAKEGICIAHSYKITSNAGE 253
 QY 250 QVGERMOCIMHNLQAQATVYVVSRRQLARVFESVVLNLTGK--VWVASEANALSRH 308
 Db 254 QSFDFKLLKLRSHLPK--ARVVAACEGEMTVGLMLAMRRLAGFEILLGSDGMADRYD 311
 QY 309 IGVGVGIORIGVNLGVALQKRAVPGIKAFEEVYARADKRAPRCHKGSCSSNOJOREQ 368
 Db 312 VV--DGYQK-EAVGGITTLQSS--PDVKPDDYLL--KLRPETNLRNPV-----FQ 355
 QY 369 AFMAHTMP--KLKAFSM--SSAYN-----AVYAAVAVAHGLH--QL 403
 Db 356 EFWQIRFOCRLEGFQAEQENKYNKTCNSLTLRTHYVDSKMGFVNAITYSMAGLHNMOM 415
 QY 404 LGCA--SGACSR-----GHVYPMOLLQIHKVHFL--LAKDTVAFNDNRDPLSSYNIAND 455
 Db 416 SLCPGYAGLCDAKMRPIDR-----KLDSLMKTNTGVSQDMILLFDENGDSPPRYELIMNK 471
 QY 456 WNGPKWTFVYLVGSSTVPVQVLINETKTIQWKGKNOVPRKSVSSDCLSEHQRYV--TGFPH 514


```

Db 137 TGVIGSYSDVSIQVANLLRFLQIPQISYASTSAKLSDKSRDYFARTVPPDFFOAKAMA 196
Qy 201 LLLQKFGMTWISLVSSDDYQGLQVQALQENATGOGICIAKNDIMPSSAQVDERMOCLM 260
Db 197 EILFFEMWTYVSTVASBEDYGETEIEAFELARARNICVA-----TSEKGRAMSRAAF 250
Qy 261 RHLQAQ-----GATVVVVFSSRQLARVFEFESVLTNLTKWVASEAMALSRHITGVGCI 315
Db 251 EGVYRALLQKPSARVAVLFTSEDAARELLATQRLN-ASFTWVASDGMGLESTV----- 303
Qy 316 QRIQVGLVALQKRAVPGKAEEEA-YARAD-----KRAPRCHKGSW-----C 358
Db 304 -----VAGSERAAAGAITIELASYIPISDFASYQSLDMNNSSRNPMWFREFEERPHC 355
Qy 359 SSNQLCRECAFMHTPKLAKAFSSMSAYNAVYAVANHLLHOLLG-----CASG 409
Db 356 SFPR--KDCAAHSLRAVFEFESKIMFVNV--AVYAAHALLHMHRRALCPNTHLCDAM 410
Qy 410 ACSGRVYPMQLERQIKHVEHLH-----KDTVAFNDNRDPLSSYNTIAMDMN 457
Db 411 RPNVGR-----RLTK-DEVLNVKFDAPRPADTDDEVDFRFGDQIGKYNIFTYLR 461
Qy 458 GP-KWTFYVAGSTWSPVQVQINETKIQWNGKD-NQVPKSVSSDCLGKHQVVTGFPHC 515
Db 462 GSGRYRQKVQ--YMAE-GILDTSFIPWASPSAGPLPASCSEPCLONEKVSQVPGEV 518
Qy 516 CFECVPCGAGTFELKSDLYRCOPCGKEWAPESQTF--CFRTVVFALREHTSVL--LA 572
Db 519 CMLCTIPQOPEY--RLDEFTADCGLGTV--PNASLTGCFELPQET--TRMDAMAVGPT 573
Qy 573 ANTLLELLLLTAGLTFAMHLDTPVYRSAGRLCFMLGSLAAGSGLYGFGEPTPAC 632
Db 574 IACGALATLEVLVGFVRHNAFTPVKASGRRELYILGGLVLCYCMFVFIAKSTANCT 633
Qy 633 LRQALFALGFTIFISCLTVNSFOLITF--KSTKYPTFTHANVQNHGAGLFWISSAA 689
Db 634 LRRIGLTAFSVCSYALTTKTRIRARIFGARREGAQRPRF-----ISPAS 678
Qy 690 QLLTCL-----TWLVVMP-----LPAREYGRPHLVMLBCT--ETNSLIGFIL 730
Db 679 QVALCCLALISQGLIIVAMLVVEARPGKETAPER--EVTYTLCKNHDSASMLG-- 730
Qy 731 AFLVNGLLISAFACSYLGDLPENYNEAKCVTFSLFNFVSIATFTTASVYDGKYLPA 790
Db 731 SLAVNVLILACTLYAFKTRCPENNEAKFIFGFTWYTCIIMLAFPIFYVTSDDYVQ 790
Qy 791 ANMAGLSSLSG--FGGYFLPKCYVILCRPDLNSTEHFQASIDYTRCGS 840
Db 791 TTTMCVSYLSGSVVLGCLFAPKHLIILFOPKRVVSH-----RAPTSRFGS 837

RESULT 13
JH0562
metabotropic glutamate receptor 3 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence-revision 30-Jun-1992 #text-change 13-Sep-1998
C:Accession: JH0562
R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A:Title: A family of metabotropic glutamate receptors.
A:Reference number: JH0561; MUID:92110002
A:Accession: JH0562
A:Molecule type: mRNA
A:Residues: 1-879 <TAN>
A:Experimental source: Brain
C:Comment: This protein is coupled to a G protein and evokes a variety of functions by m
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F:1-22/Domain: signal sequence #status predicted <sig>
F:23-879/Product: metabotropic glutamate receptor 3 #status predicted <MET>
F:577-599/Domain: transmembrane #status predicted <TRI>
F:614-634/Domain: transmembrane #status predicted <TII>
F:646-664/Domain: transmembrane #status predicted <III>

```

```

F:689-709/Domain: transmembrane #status predicted <TIY>
F:735-756/Domain: transmembrane #status predicted <TRV>
F:770-791/Domain: transmembrane #status predicted <TRI>
F:804-828/Domain: transmembrane #status predicted <TII>
F:209,292,414,439/Binding site: carbohydrate (asn) (covalent) #status predicted
F:610,845/Binding site: phosphate (Ser) (covalent) #status predicted

```

Query Match 15.3%; Score 688; DB 2; Length 879;

Best Local Similarity 25.8%; Pred. No. 2,5e-44;

Matches 227; Conservative 139; Mismatches 355; Indels 160; Gaps 35;

```

Qy 31 DFTLPGLVILGLTPRLHSGCLQVRHREPVTLCDRSCSFNE-HGYHLFOAMRLGVEINNS 89
Db 32 EIKIEGDLVGLGRLPINE-----KGTETECGRINERGRIGRLEAMFALDEINKD 82
Qy 90 TALLPNTILGYQLDYVCSANVYATIRVLSL-----PGQHIELOGLD 133
Db 83 NYLLPGVYKLVGHILDTG--SRDTYALBQSLFVFRASLTKEVDEAEYMCPSGAYAIQENIPL 140
Qy 134 LHSPTVLAIVGPDSTNAATTAALLSPFLVPMISYASSSETISYKRPSTFLTPMDK 193
Db 141 L-----IAGVIGGSYSSYIQVANLLRFLQIPQISYASTSAKLSDKSYDYFARTVPPDF 195
Qy 194 YQVETVNLLOKFGMTWISLVSSDDYQGLQVQALQENATGOGICIAFKDIMPSAQQVD 253
Db 196 YQAKAMAILFPMWTYVSTVASBEDYGETEIEAFQGRARLRNICIATAE-----KYGR 249
Qy 254 ERMO-----CLMRHLAQ-AGATVVVVFSSRQLARVFEFESVLTNLTKWVASEAMALSRH 308
Db 250 SNIRKSYDVIRELLQKPNARVVVLFMRSDSRRLIAANRVN-ASFVWVASDGMGQAES 308
Qy 309 ITGPGIQRIGMIVGVALQKRAVPGKAEEEAARADAPRCHKGSW-----C 358
Db 309 I--VKGSEHV--AVGATLELASHVROFDRIFQSLN--PYNHNRPFEDFPEQKFOC 361
Qy 359 S-----SNQLCRECAFMHTMPKLFKAFSSMSAYNAVYAVANHLLHQLQ-----LLGAS 408
Db 362 SLQKRNHRQYCKHLALIDSNNYEQ-----ESKIMFVNAVYAMAHKMQRLCPRTT 416
Qy 409 GACSRGRVYPMQL--EQLHVEHLHMDIYAFNDN-----DPLSSNTIAM 454
Db 417 KLCDAKMLTDKKLYKELKLINE-----TAPFNKKGADSIKFDTPGDMGRVNVNL 471
Qy 455 DMNPGKTFYVIGSSTWSPVQVQINETKIQWNGKDNOVPKSVSSDCLGKHQVVTGFPH 514
Db 472 QQTGKTSYLVGH--WAEF--LSLDVDSIH--SRNSVPTSCSDPCAPNEMKMQPQDV 526
Qy 515 CFECVPCGAGTFELKSDLYRCOPCGKEWAPESQTF--PRVVFALREHTSW----- 568
Db 527 CWCICIPCEPEYEL--VDEFTCMDCGPGQWPTADLSGCTNLPEDYI-----KMEDAMALGP 580
Qy 569 VLLAANTLLLLTAGLTFAMHLDTPVYRSAGRLCFMLGSLAAGSGLYGFGEPTR 628
Db 581 VTIAICIGLTCIYIVT--VFTKHNTPLVKASGRRELYILGGLVLCYCMFVFIAKSP 638
Qy 629 PACILRQALFALGFTIFISCLTVNSFOLITF--KSTKYPTFTHANVQNHGAGLFWMI 685
Db 639 VICALRRIGLCTSPATCYSAITLTNCTARIFDGVKCAQQRKE-----I 683
Qy 686 SSAQDLTCL-----TWLVVMP-----LPAREYGRPHLVMLBCTETNSIGF 728
Db 684 SPSSGVFICLGLIIVQIYMWVSWLILFPGTRRYTLPEKR-----EYILKCNKWS-SM 737
Qy 729 ILAFLVNGLLISAFACSYLGDLPENYNEAKCVTFSLFNFVSIATFTTASVYDGKYL 788
Db 738 LISLIVDVAVILCTVYAFKTRCPENNEAKFIFGFTWYTCIIMLAFPIFYVTSDDY 797
Qy 789 PAANMAGLSSLSG--GGYFLPKCYVILCRPDLNSTEH 826
Db 798 VQTTMC--ISVLSGFEVVLGCLFAPKHLIILFOPKRVVSH 837

```



```

0Y 389 ---AURVAVANGHQ---LLGASAGC-----SRGVYPMQLEOIHAKYHLLHKD 435
Db 394 KQVQYIDAVYAIINALISMHOALCPGHTGICLPRAMEPDDGT---LHATRAVRFGNSAG 449
QY 436 T-VAENDNRDELSSYNIAND-WNGPKWTFYVIGSSWSPVOLINSTRKIOMHGKDQVP 4933
Db 450 TPVAMENEGDAPGRYDIFQYQATNGSSASSGGYQAVQWAEALRLDEVLRWMSGDDPHEVP 5088
QY 494 KSVSSDCLBEHORVYVGHHCCEYCPCCAGFTLANKSDLYRQPCCKEMAEBSQTCF 5533
Db 509 PQSCSLPCGPBEKKMKYKVPCCWCHCACOGYR--QYDFETCEACGDMRPPNHTGCR 5666
QY 554 PRTVFTALREHTSMVLIANTLLILLLLG-----TAGFEAMHLDPPRVASGRCLFL 6078
Db 567 PTYPVRL-----TWSSPMALPRLVLAVLGIMATTTIMATFMNRNDPIYRASGRLESTYV 6200
QY 608 MUGSLAAGSGSLYCFEGEPTRPACLLRQALFALGFTFTLSCLYVRSQOLIIF---KEST 6644
Db 621 LLTIGIFLIYATFTFMAVEPCALICQARLLLGIGTTLTSYALLTKTNRIYRFEQGRSV 6860
QY 665 KVPFFYHAMVQNHGAGLFVMISSAOLLI-----CLTWLVWYPTLPAREYOR-- 7111
Db 681 TPRPF-----ISPSQVLYITPGLTSLQVGYIAMLGAPQPHSYVIDIEQR 7255
QY 712 --FPHLV--MLECTETNSLGIIFLAFYLVNGLTISAFACSYLGDNDPENYNEACVYTESL 7677
Db 726 TVDEQAGQVYLC-DMSDDLSTIGCLGSLTLLMTYCAIYALRKARVPETPFNEAKPIGFTMY 7844
QY 768 FNFPMWIA----FPTTASVYDQKTLPRANMAGLS--SLSGSFGFYLPKCYVILCRDNLN 8222
Db 785 TTCTIIMIAEVPFEGTQASAKIYIQTTTLTAVSLTSSLASVSLGVLVPKYVILLFHEQN 8444

```

Search completed: June 11, 2002, 10:19:35
Job time: 2829 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 09:19:53 ; Search time 24.84 Seconds

(Without alignments)
826,970 Million cell updates/sec

Title: US-09-819-946-2

Perfect score: 4493
Sequence: 1 MLCTARLVGLQLISCCWA.....NSTEHPQASIDYTRCGST 841

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 segs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued_Patents_AA:*
1: /cgn2.6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/prodata/1/1aa/PCFUS_COMB.pep:*
6: /cgn2.6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1142	25.4	1078	1	US-08-485-588-7 Sequence 7, Appl1
2	1142	25.4	1078	1	US-08-484-565-7 Sequence 7, Appl1
3	1142	25.4	1078	2	US-08-480-751-7 Sequence 7, Appl1
4	1142	25.4	1078	2	US-08-943-986-7 Sequence 7, Appl1
5	1142	25.4	1078	3	US-08-353-784-7 Sequence 7, Appl1
6	1142	25.4	1078	3	US-08-484-719B-7 Sequence 7, Appl1
7	1142	25.4	1078	4	US-08-484-159-7 Sequence 7, Appl1
8	1130	25.2	1088	1	US-08-485-588-6 Sequence 6, Appl1
9	1130	25.2	1088	1	US-08-484-565-6 Sequence 6, Appl1
10	1130	25.2	1088	2	US-08-480-751-6 Sequence 6, Appl1
11	1130	25.2	1088	2	US-08-943-986-6 Sequence 6, Appl1
12	1130	25.2	1088	3	US-08-353-784-6 Sequence 6, Appl1
13	1130	25.2	1088	3	US-08-484-719B-6 Sequence 6, Appl1
14	1130	25.2	1088	4	US-08-484-159-6 Sequence 6, Appl1
15	1127.5	25.1	1085	1	US-08-485-588-5 Sequence 5, Appl1
16	1126.5	25.1	1085	1	US-08-484-565-5 Sequence 5, Appl1
17	1126.5	25.1	1085	2	US-08-480-751-5 Sequence 5, Appl1
18	1126.5	25.1	1085	2	US-08-943-986-5 Sequence 5, Appl1
19	1126.5	25.1	1085	3	US-08-353-784-5 Sequence 5, Appl1
20	1126.5	25.1	1085	3	US-08-484-719B-5 Sequence 5, Appl1
21	1126.5	25.1	1085	4	US-08-484-159-5 Sequence 5, Appl1
22	1126.5	25.1	1085	4	US-08-485-588-8 Sequence 8, Appl1
23	1123	25.0	1079	1	US-08-480-751-8 Sequence 8, Appl1
24	1123	25.0	1079	2	US-08-480-751-8 Sequence 8, Appl1
25	1123	25.0	1079	2	US-08-943-986-8 Sequence 8, Appl1
26	1123	25.0	1079	3	US-08-353-784-8 Sequence 8, Appl1
27	1123	25.0	1079	3	US-08-484-719B-8 Sequence 8, Appl1

28	1123	25.0	1079	3	US-08-484-719B-8 Sequence 8, Appl1
29	1123	25.0	1079	4	US-08-484-159-8 Sequence 8, Appl1
30	1093.5	24.3	1027	4	US-09-162-021B-2 Sequence 2, Appl1
31	1088	24.2	1219	2	US-08-687-289A-6 Sequence 6, Appl1
32	777	17.3	1058	2	US-08-687-289A-5 Sequence 5, Appl1
33	741.5	16.5	1180	4	US-08-660-148-2 Sequence 2, Appl1
34	741.5	16.5	1212	4	US-08-660-148-5 Sequence 5, Appl1
35	733	16.3	1056	2	US-08-687-289A-7 Sequence 7, Appl1
36	727	16.2	906	5	PCT-US91-09422-17 Sequence 17, Appl1
37	727	16.2	1199	1	US-08-041-338-2 Sequence 2, Appl1
38	727	16.2	1199	1	US-08-463-642-2 Sequence 2, Appl1
39	727	16.2	1199	1	US-08-455-602-2 Sequence 2, Appl1
40	727	16.2	1199	2	US-08-465-157-2 Sequence 2, Appl1
41	727	16.2	1199	5	PCT-US91-09422-2 Sequence 2, Appl1
42	724	16.1	906	1	US-08-486-270-2 Sequence 2, Appl1
43	724	16.1	906	3	US-08-367-264-2 Sequence 2, Appl1
44	724	16.1	1056	2	US-08-687-289A-8 Sequence 8, Appl1
45	720.5	16.0	1180	1	US-08-486-270-8 Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-08-485-588-7
Sequence 7, Application US/08485588
Patent No. 5688938
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451

ELLING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38, 179
REFERENCE/DOCKET NUMBER: 213/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-588-7

Query Match 25.4%; Score 1142; DB 1; Length 1078;
Best Local Similarity 31.0%; Pred. No. 2e-105;
Matches 276; Conservative 163; Mismatches 358; Indels 92; Gaps 22;

QY 16 SCW---AFACHSTESSPD--FTLPDYLALGLPLHSGCL----QVRRPEVTLCDRGC 66
DB 5 SCCWVLLALTWHTSAVGPQRAKKGDITLGLFPIHFVGAADKDDLSKSPSEVCEIR-- 62
QY 67 SENEGHYLFQAMRLGVEEINNSTALLPNTLGYQLDYVCSANVATLRVLSLPGOH 126
DB 63 -YNFGRFRLQAMFAIEEINSSPALPLTLGRIYFDICNTVSK--ALEATLSFAQAK 119
QY 127 IELQG----DLHSPVLAIVIGDPSTNRAATTAALSPFLVPMISVASSSETLSVKRO 181
DB 120 ISLNIDFCNCSEHPISTAVYATGSGSTAVANLGLFYTPQVYASSRLSNKNQ 179
QY 182 YPSFLTITNDKQYQVETWLLQKRGWTISLVGSSDDYQGLQVQLAEMQATGQICIAF 241
DB 180 FSPFLTIPNDHQATAMADILEFRMNWVGTIAADDDYGRPGIEKFRFEAEERDCIDF 239
QY 242 KDIMPESAQVDERMOCLRHLAGATVAVVSSRQLARVEFESVLTNLGKVVASE 301
DB 240 SLLI--SQSDEEIQHVEVIONSTAKIYVSSGPDLEPLIKETVRNITGKITLASE 297
QY 302 AVALSRHITGVPQIGRICVGLVATOKRAVPGIKAF-----EAYARADKEAPP 351
DB 298 AAASSSLAMPQYFHVYGGTIGFALKAQIPGFRFLKHYHPRKSVHNGFAKEFMEETEN 357
QY 352 CH-----KG-----SMCSSNQLCREQAFAMHTPKLAFMSMSA 386
DB 358 CHLOEGAKGPLVDVTFRLGHESGDRFSNSTAFRPLCTGDENISSVETPYIDYTHLRIS 417
QY 387 YNAAYRAVVAHGLHQLLGC-----ASGACSR--GRVYPMQLLEQIHKVHFLHK-DTV 437
DB 418 YVYLAIVASIALADQIYCLPGRGLFTNGSCADIKKVAQVQLKRLRLNFTNNNGEV 477
QY 438 AENDNRDPLSSNITIAMDMNGPKWFTTVLGSSWSVPY-----OLNINETKIOMH 486
DB 478 TDECGDLVANGSIINMHLSPED-----GSIVFKEVGYNYAKKGERLFTNEEKILMS 531
QY 487 GRDNOYKVCSSDLEQHQR--VVTGFHHCCEVCYPCGAGTFLEKSDLRQPCGSEEMA 545
DB 532 GFSREVPFSNCRDLGTRKGIIESEPCCCEVCPCPGEGYSDDETASACNCPDPEWS 591
QY 546 PEGSOTCFPRVTVFLALREHTSMVLLAANTLLLLLAGTAGFAHMLDTPVYVSAGRLC 605
DB 592 NENHNSCIAKEIEFLSWTEPFGLTLFLAVLGIPLAFVLGFIKRNRPYIKATRELS 651
QY 606 FLMLGSLAAGSSLYGFCEPFRPACLFQALFALGFTTFLSLCYVRSQQLIITFEFSK 665
DB 652 YLLPSSLCCFSSSFFETIEPQDWTCRLRQPAFGISFVLCISCIIVKTRVLLVFE--AK 709
QY 666 VPTFVH-AWVQNHGAGLFVMISSAQLLCLMVLVWTFPLPARVYRPHIWMLETEIN 724
DB 710 IFTSHRRKWWGILNQLFLVFLCTFMQIVICVIMLYTAPPSSTRNOLEDELIITFCHES 769

QY 725 --SLGFIALFNLGLISAFACSYLGDLPENYNEAKCVTSLLFNFVSWIAFFT--AS 781
DB 770 LNALGFLIG--YTCLLAALICEFFAFKSRKLPEFNFAKFTFSMLFFIWMISFIPAYAS 827
QY 782 VYDKYLPANMMAGLSSSGGGYFLPKCYVILLCRPLDNSTENHQAS 830
DB 828 TY-GKFSVAVEVYALIAASFGLLACIFENKIIYILFKPSRNTIEVRCSS 875

RESULT 2
US-08-484-565-7
Sequence 7, Application US/08484565
Patent No. 5763569
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484, 565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38, 179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear


```

0Y 7 SENEGYLLPAMRGVGEVINNNPALLPNLITGOLQVDCSDSANVATLRVLSLPQOH 126
Db 63 -YNRGFWLQDAMIFALEINSSPALLPNLITGRIDTCNYSK--ALBATTISFAQOK 119
0Y 127 TELG-----DLHYSPTVLAVIGPDSNRAATTAALLSPFLVPMISYASSSETLSYKRO 181
Db 120 IDSINLDEFNCSEHIFSTIAVVGATGSVATAVANLLGLFIFYQVYASSRLLSKNQ 179
0Y 182 YPSFLRTIPNKKQVETMVLNLLQFGWTWISLVSSSDDYQOLGVQALENATGOGICIAF 241
Db 180 FKSEFLRTIPNDEHQTATAADLIEEFERNWVGTTAADDYGRPGIEKREAEEREDIDF 239
0Y 242 KDIPEFSAQVODERMOCMLRHQAQATVVVVFESSROLARFEFSEVYLTWTKGVWVASE 301
Db 240 SELT--SGYSEDEELQHVEYIOWSTAKVIVFESSGDEBLEIKELIYRNITGKIWLASE 297
0Y 302 AMALSRHTTGVPGIQRIGMWLGAALOKRAVPGKAF-----BEAVARADKEARPR 351
Db 298 AMASSSLIAMPQYFHVVGTTGFGALKAGQIEFREFELKVKHPRKSVYNGCAPEWETFN 357
0Y 352 CH-----KG-----SMCSSNOICRECOAFMATHMKIKAFSSSA 386
Db 358 CHLOGAGGLPVPDTFLRGHEESGDRFSSNSTARPLCTDENISVETPYIDYTHLRIS 417
0Y 387 YNARYAVVAHAGELHOLIGC-----ASGASR-GHUYPMOLLEOYHKKFLHK-DYV 437
Db 418 YNVLYAVYSINHALODIYTCLPGRGLTNGSCADIKKVEAMOVUKLHLNFTNNMOEY 477
0Y 438 AFNDNRDPLSSYNIITAMDNGPKMTFTVLSSWSPV-----QINETKIQWH 486
Db 478 TFEDEGGLVGNYSIIINHLSPED-----GSIFYKEVGYNYVAKKERFELINDEKILWS 531
0Y 487 GKDNQVPKSVSSOCLEHON-VYTGFGHHCCEVPCVAGCFELNKSLYKCOQCKEMEA 545
Db 532 GFSREVPKSNCSRODLGTRKGIITEGPTCCFEVCEPDEYSDETDAKCNKCPDEWMS 591
0Y 546 PEGSOCPRTVTEALAREHSWLLAANTLLLLLLLTGTAFLAMHIDTPVVRSAAGRLC 605
Db 592 NENHTSOIAKEIEFLSMTEPGIALTLFVAVLGIFLAFVGVFIKENPTIYVATINELS 651
0Y 606 FLAMGSLAAGSGSLXGFFGETPRACILRQALFALGTTIFLSCILYVSPOLIIIFKESTK 665
Db 652 YLLFSLIACPFSSSLFEGEPDQTCRLRQAPFISIVLCSICILVYTNVLLVFE--AK 709
0Y 666 VPTFYH-AMVONHGAGLEVMVSSAOLILCTCTWVVMVPLPAREYORPRLHMECETN 724
Db 710 IPTSHRKRMGLNLOFLVFLCTMQOYIYCIWLYTPARPSRYKQOLEDELIFITCHEGS 769
0Y 725 -SIGFLIAFLYNGLLSISAFAGSYLGKDRPENYNEAKCVYFSLLEFNFSWIAFEFTT-AS 791
Db 770 LMAIGFLIG--YTCILAAICFPFAFKSRNKPENENAEKFTISMLFIIWISFIPAYAS 827
0Y 782 VYDGKYLPAAMMAGLSSSLSGPGGYFLPKCYIYLCRPDLNSTPHRAS 830
Db 828 TY-GEFSAVEVIALIASFGLIACIFENKIYLLIFPDSRNTLEEVCS 875

```

RESULT 4
US-08-943-986-7
; Sequence 7, Application US/08943986
; Patent No. 5962314
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street

```

1 CITY: Los Angeles
2 STATE: California
3 COUNTRY: USA
4 ZIP: 90071
5
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: FASTSEQ
11
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/08/943,986
14 FILING DATE: 03-OCT-1997
15 CLASSIFICATION: 530
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: 08/484,565
18 FILING DATE: 7-June-1995
19 APPLICATION NUMBER: 08/353,784
20 FILING DATE: 9 December, 1994
21 APPLICATION NUMBER: PCT/US/94/12117
22 FILING DATE: 21 October, 1994
23 APPLICATION NUMBER: U.S. 08/292,827
24 FILING DATE: 23 August, 1994
25 APPLICATION NUMBER: U.S. 08/141,248
26 FILING DATE: 22 October, 1993
27 APPLICATION NUMBER: U.S. 08/009,389
28 FILING DATE: 23 February, 1993
29 APPLICATION NUMBER: U.S. 08/017,127
30 FILING DATE: 12 February, 1993
31 APPLICATION NUMBER: U.S. 07/934,161
32 FILING DATE: 21 August, 1992
33 APPLICATION NUMBER: U.S. 07/834,044
34 FILING DATE: 11 February, 1992
35 APPLICATION NUMBER: U.S. 07/749,451
36 FILING DATE: 23 August, 1991
37 ATTORNEY/AGENT INFORMATION:
38 NAME: Heber, Sheldon O.
39 REGISTRATION NUMBER: 38,179
40 REFERENCE/DOCKET NUMBER: 213/006
41 TELECOMMUNICATION INFORMATION:
42 TELEPHONE: (213) 489-1600
43 TELEFAX: (213) 955-0440
44 TELEX: 67-3510
45
46 INFORMATION FOR SEQ ID NO: 7:
47 SEQUENCE CHARACTERISTICS:
48 LENGTH: 1078 amino acids
49 TYPE: amino acid
50 TOPOLOGY: linear
51 MOLECULE TYPE: protein
52
53 US-08-943-986-7
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912

```

Db 240 SELI--SQYDEEIOHVEVIONSTAKVIYVSSGPDLEPIKEIVRNITGKIWLASE 297
QY 302 AMALSRHITGVPGIORIGMVGVVALOKRAVPGIAF-----EAYARADKEARPP 351
Db 298 AMASSSLIAMQYFHVVGTTIGFALKAGQIPGFRFLKKYHPRKSVANGFAKEMEETFPN 357
QY 352 CH-----KG-----SWCSSNOLCRECOAFMAHTMPKIAKAFSSMSA 386
Db 358 CHLOGAKGRLPVDTPLRGHESGDRFNSSTAFAPLCTGDBENISSVETPIDYTHLRIS 417
QY 387 YNAVRAVVAHGLHQLG-----ASGACSR-GRAYPMOLLEQIKHVEHLK-DTV 437
Db 418 YNVLAVYSIAHALQDIYTCPLRGRLFTNGSCADIKKYEAMQVLYKHLRLNFTNMNGQV 477
QY 438 AFNPNRDLSSYNTIAMDMNGPKMTFTVLSSTMSPV-----OLNINENKIOH 486
Db 478 TFDGCGDLVGNYSIINHLSPED-----GSIVREKGYVAVYAKKGRFLINEKILMS 531
QY 487 GKDNQVPRKSVCSDDLSEHOR-VYTGPHHCCFECVPCGAGTFLNKSLLYRCOPCGKEEMA 545
Db 532 GFSREVPFNSNRDCLAGTRKGIIEGEPTECECECPDGEISDETDASACNKCDDPMS 591
QY 546 PEGSOTCEPRVIVLALREHTSWVLLAANTLLLLLTAGLPAHMLDTPVYRAGGRLC 605
Db 592 NENFTSCIAKEIEFLSWTEPFGIALTLFVAVGIFLTAFLVETIKFRMTPIYKATNRLS 651
QY 606 FLMGSLAAGSGSLYGFGEPTRACLIRALFALGFTFLSCLTVRSFOLILIEFKSTK 665
Db 652 YLLPFLSLCCSSSLFEGEQDMTCRLRPAFGISFLVCLISCLIVKRNRLVFE--AK 709
QY 666 VPFRYH-AMVONHAGLFVMISSAOLILICLWLVTLPAREXORPPLHMECTETN 724
Db 710 IPIFSHRRKMWGLNQFLVLCFEMQIYCIWLYTAPSSYRNGLEDEILFIITCHGS 769
QY 725 --SLGFLIATLYNGLTISAFACSYLGKDPENINEAKCVTFSLLENVSIATFTT-AS 781
Db 770 LMAIGFLIG--YTCLLAICFFFAKSRKLEPENFEAKFITFSMLIFIVMISFIPAYAS 827
QY 782 VYDKKYLPAANMAGLSSSGFGYFLPKCYVILCRPDLNTEHQS 830
Db 828 TY-GKEVSAVEVIAIILASFGLLACIFPNKIYILFKPSRMTIEVRCS 875

RESULT 5
US-08-353-784-7
Sequence 7, Application US/08353784
Patent No. 6011068
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert.
APPLICANT: Bradford C. Van Wagenen, Manuel
APPLICANT: F. Balandrin, Forrest H. Fuller,
APPLICANT: Eric G. Delmar, and Scott T. Moe
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,784
FILING DATE: 9 December, 1994

CLASSIFICATION: 514
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 8
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,787
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 209/069
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-784-7

Query Match 25.4%; Score 1142; DB 3; Length 1078;
Best Local Similarity 31.0%; Pred. No. 2e-105;
Matches 276; Conservative 163; Mismatches 356; Indels 92; Gaps 22;

QY 16 SCCW---AFACHSTSPD--FTLPDYLLAGLFFLHSGCL---QVRRHPEVTLCDRSC 66
Db 5 SCCWLLALTMHTSAVGPDRQAKKGDILGLFPIHGVAAKDDKSRPESVECR-- 62
QY 67 SFNEGHYLFQAMRLGVEEINNSTALLPNTLYCYLDVCSANAYATLRLVSLPGQH 126
Db 63 -YNRFGFRLQAMIRAIIEINSSPALLNLTLYGRIFDTCNTVSR--ALEATLSEVAQNK 119
QY 127 IELQ-----DLHSPYVLANYIGDSTNRATTAALLSPFLVPMISVAASSETISVKRQ 181
Db 120 IDSLNLDRCNCEHPTSTIAVAVGATSGVSTAVANLGLFYIPQVSTVASSRLLSNKQ 179
QY 182 YPSFLRITPNDRKYQVETWVLLLOKFGWTWISLVGSSDDYGGOLGVDALENOATGOGICAF 241
Db 180 FKSFLRTIPNDHQATMAADITIEFRMNMVGTIADDDVGRGIGIKFREAREERICIDF 239
QY 242 KDIMPFAQVDERMOCIMRLHLAGATVYVVFSSRQLARVEFESVLTNLTKYVAVASE 301
Db 240 SELI--SQYDEEIOHVEVIONSTAKVIYVSSGPDLEPIKEIVRNITGKIWLASE 297
QY 302 AMALSRHITGVPGIORIGMVGVVALOKRAVPGIAF-----EAYARADKEARPP 351
Db 298 AMASSSLIAMQYFHVVGTTIGFALKAGQIPGFRFLKKYHPRKSVANGFAKEMEETFPN 357
QY 352 CH-----KG-----SWCSSNOLCRECOAFMAHTMPKIAKAFSSMSA 386
Db 358 CHLOGAKGRLPVDTPLRGHESGDRFNSSTAFAPLCTGDBENISSVETPIDYTHLRIS 417
QY 387 YNAVRAVVAHGLHQLG-----ASGACSR-GRAYPMOLLEQIKHVEHLK-DTV 437
Db 418 YNVLAVYSIAHALQDIYTCPLRGRLFTNGSCADIKKYEAMQVLYKHLRLNFTNMNGQV 477

```

QY 438 AFNNRPDLSYNNIIADMDMNPCKWTFVLGSGTSPV-----Q0INETRKHOM 486
Db 478 TFDCEGDLYGNYSIIINHLSPED-----GSIVKEGYNNVYAKKERLFINDEKILMS. 531
QY 487 GKDNQVPKSVCSDCLEGHOR-VVTGFHCCFECVPCGAGTFLNKSDDLRYBCPGCKGEWA 545
Db 532 GFSREVPFSPNSRDLCLAGTRKGIIEGBPTCCFECVECPDGEYSDETDASACNCKCPDDFWS 591
QY 546 PEGSOTCPPRVYVFLALREHISWVLNANTLILLLLIGTAGLPRAMHLDTPVVSAGGRLC 605
Db 592 NENHTSIAKIEELSWTEPPGIALTLFVYLGJFLTFVGLVGFLEKFNPTIVAKTNRELS 651
QY 606 FLIMGSLAAGSGSYLGEFGEPTPRACILROLALFGTIFLSCITVSPFOLLIFPKSTK 665
Db 652 YLLFSLTLCSSSLFETGEQDMTCRLRPABISIVYLICSLVAKTNRYLVFE--AK 709
QY 666 VPTEYH-AMVONHGAGLEFVMISSAOLLICLTWLDVVTPLPAREYORPHLVMECTETN 724
Db 710 IPTSPHRRKMGLNLQFLLVFLCTPMDIVCIWMLYTPAPSSYVRQOEDLEDEIFITCHEGS 769
QY 725 --SLGFIILAEIYNGLLISAFACYSLGKDLPEYNTNEAKCYTFSLELNFVSWIAFFT--AS 781
Db 770 LMAIGLFLIG--YTLLOLLAICFFFAFKSRKRPENENNEAKFTFSMLFIWISIFIPAYAS 827
QY 782 VYDGDYDPAAMMAGISLSSGFGYFLPKCYVILLCAPDINSTHHPAS 830
Db 828 TY-GKFSYAVEYVAILLASFCGLACIFPNKYYIIILFKPSRTTIEVACS 875

```

US-08-484 6
US-08-484-719B-7
Sequence 7, Application US/08484719B
Patent No. 6031003
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert,
APPLICANT: Bradford C. Van Wageningen,
APPLICANT: Manuel F. Balandrin,
APPLICANT: Forrest H. Fuller, Eric G.
APPLICANT: Delmar Scott T. Moe
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS Word
SOFTWARE: Fastseq for Windows Version 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,719B
FILING DATE: 7 June, 1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127

```

1      FILING DATE: 12 February, 1993
2      APPLICATION NUMBER: U.S. 07/934,161
3      FILING DATE: 21 August, 1992
4      APPLICATION NUMBER: U.S. 07/834,044
5      FILING DATE: 11 February, 1992
6      APPLICATION NUMBER: U.S. 07/749,451
7      FILING DATE: 23 August, 1991
8      ATTORNEY/AGENT INFORMATION:
9      NAME: Douglas C. Murdock
10     REGISTRATION NUMBER: 37,549
11     REFERENCE/DOCKET NUMBER: 213/007
12     TELECOMMUNICATION INFORMATION:
13     TELEPHONE: (213) 489-1600
14     TELEFAX: (213) 955-0440
15     TELEX: 67-3510
16     INFORMATION FOR SEQ ID NO: 7:
17     SEQUENCE CHARACTERISTICS:
18     LENGTH: 1078 amino acids
19     TYPE: amino acid
20     STRANDEDNESS: single
21     TOPOLOGY: linear
22     MOLECULE TYPE: protein
23     US-08-484-719B-7

```

Query Match	25.4%;	Score 1142;	DB 3;	Length 1078;
Best Local Similarity	31.0%;	Pred. No. 2e-105;		
Matches 276;	Conservative 163;	Mismatches 358;	Indels 92;	Gaps 22

```

0Y 16 SCOM---AFACHSPRESSPD---FTLPGYLLAGLEPHHSGCL---OVRHHPEVTLCDRSC 66
Db 5 SCOMVLLALHTHTSYGPDQAOAKKKDIIIGLGFPHHGYAANKQDKLSRPESECTR-- 62
0Y 67 SFNEHGHYLFQAMRLGVEEINNSTALLPNTITLGOLYDVCSDSANYATATLVLSLPGOH 126
Db 63 -YNFGEFWMLOAMIFALEEINSSPALLPNLTGLRIDPTCNVSK--ALBATSIFYAOK 119
0Y 127 IELOG----DLHNSPVILAVIPDPSNRRATTAALISPIVMISSAASETILSKRQ 181
Db 120 IDSILNDEFCCSCSEHPISTIANVAVATSGYSTAVANILGLFYLIOVSYASSRLLSKNQ 179
0Y 182 YPSFLRTIPNDKRYOVETVLLQKFGMTWISLVGSSDDYDQGLQVQALENOATGQGLCIAF 241
Db 180 FKSFRLTIPNDEHQATAMADIIIEYFRNMWVGTIAADDDYGRGIEKPREAEERDIDCF 239
0Y 242 KDIMEFSQVODERMOCLMRLLAQAQATVYVYESSRQARVFEFSVYLTNLTKGVWASE 301
Db 240 SELI--SOYSDEEIOHVEVIONSTKAVIYFESSGPDLEPLKEIYRNNTIGTWIASE 297
0Y 302 AMALSRIHTTGPVGIORIGMVGVAVIQKRAVAGLAF-----EBAVARADKEAPR 351
Db 298 AMASSSILAMQYFHVYVGGTIGFALKKQPIDGFEFLKYHPRKSVYNGAKKEWETFN 357
0Y 352 CH----KG-----SMCSSNOJCRECOAFMAHTMFKLAFSSSSA 366
Db 358 CHLOEGAGGPIPVDTFLRGHEESDREFSNSSTARPLCTGGENISSVETPEYIDYTHLRIS 417
0Y 387 YNAYRAVYAAVHGLHOLLG-----ASGASR-GGVYPMOLLQJHKKHYLLHK-DIV 437
Db 418 YNVYIAVYSIAHAIODITYTCIPGSGLEFTNOSCAIKVEAMOVYLLKLRHLNFTNNNGEOV 477
0Y 438 AFNDNRDPLSSYNIILADMDNPKMTFVLVGSSTWSPV-----OLNIMETKIOHM 486
Db 478 TFEDEGDDLVGNYSIINHLSPED-----GSIVKEGYNYVYAKKGERFLINEKTIWS 531
0Y 487 GKDNQVPKSVCCSDCLEGHOR-VVTGFHNCCEFCVPCGAGTFLNKSDLYRCQPCGKEEMA 545
Db 532 GFSREVPFNSRCLAGTRKGIIEGEPGCCFECEVCPDESEDETASACNKCPRDFWS 591
0Y 546 PEGSOTCEPRVYVLLALREHNSWVLLAANTLILLLLTAGLGRAMHLDTPVYVSAGGRIC 605
Db 592 NENHTSIAKIEFLSWTEPGIALTLEAVYGLTFAVGLFKENRNTIYVATNRELS 651

```

```

01 606 FLMGLSLGAGSGSLYGFGEPTBRACLLRQALFALGFTIFLSCLTVASFOLIIIFKRSK 6655
02 652 YLLFLSLCCSSSLFTIGFGEQDMTCRHRPAPFISIVLCISCLIVATNRVLVE--AK 70909
03 666 VPTFYH-AMVONHGAGLEFVMISSAOLLICLTWLVTWTPPLPAREYQRPPLVLMJECETN 7244
04 725 --SLGFIILAFILNGLLISAPACSYLKGKDLPENYNEAKCYTFLSLFNFVSMIAFTT-AS 7811
05 770 LMAIGFLIG--YTCLLLAICFFFAFKSRKLPENYNEAKFTFSPMLIFFIYISIPAYAS 8271
06 782 YVDGKYLDAANMMAGLISLSGFCGYFLPKCYVILICPDLNSTPHQAS 830
07 828 TY-GKFSVAVEYIALIASFELLACIFPNKKIYIILFPRSRNITEVNC 875
08
09 RESULT 7
10 US-08-484-159-7
11 ; Sequence 7, Application US/08484159
12 ; Patent No. 6313146
13
14 GENERAL INFORMATION:
15 APPLICANT: Bradford C. Van Wagenen
16 APPLICANT: Manuel F. Balandrin
17 APPLICANT: Eric G. Del Mar
18 APPLICANT: Edward F. Nemeth
19 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
20 MOLECULES
21 NUMBER OF SEQUENCES: 20
22
23 CORRESPONDENCE ADDRESS:
24 ADDRESSEE: Lyon & Lyon
25 STREET: First Interstate World Center
26 STREET: Suite 4700
27 STREET: 633 West Fifth Street
28 CITY: Los Angeles
29 STATE: California
30 COUNTRY: USA
31 ZIP: 90071
32
33 COMPUTER READABLE FORM:
34 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
35 COMPUTER: IBM PC compatible
36 OPERATING SYSTEM: PC-DOS/MS-DOS
37 SOFTWARE: FASTSEQ
38
39 CURRENT APPLICATION DATA:
40 APPLICATION NUMBER: US/08/484,159
41 FILING DATE: 7 June, 1995
42 CLASSIFICATION: 435
43
44 PRIOR APPLICATION DATA:
45 PRIOR APPLICATION DATA: including application
46 PRIOR APPLICATION DATA: described below: 9
47 APPLICATION NUMBER: 08/353,784
48 FILING DATE: 9 December, 1994
49 APPLICATION NUMBER: PCT/US/94/12117
50 FILING DATE: 21 October, 1994
51 APPLICATION NUMBER: U.S. 08/292,827
52 FILING DATE: 23 August, 1994
53 APPLICATION NUMBER: U.S. 08/141,248
54 FILING DATE: 22 October, 1993
55 APPLICATION NUMBER: U.S. 08/009,389
56 FILING DATE: 23 February, 1993
57 APPLICATION NUMBER: U.S. 08/017,127
58 FILING DATE: 12 February, 1993
59 APPLICATION NUMBER: U.S. 07/934,161
60 FILING DATE: 21 August, 1992
61 APPLICATION NUMBER: U.S. 07/834,044
62 FILING DATE: 11 February, 1992
63 APPLICATION NUMBER: U.S. 07/749,451
64 FILING DATE: 23 August, 1991
65 ATTORNEY/AGENT INFORMATION:
66 NAME: Heber, Sheldon O
67 REGISTRATION NUMBER: 36,119
68 REFERENCE/DOCKET NUMBER: 214/101
69 TELECOMMUNICATION INFORMATION:

```

[illegible]

TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION DATA: including application
PRIORITY APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1088 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-565-6

Query Match 25.2%; Score 1130; DB 1; Length 1088;
Best Local Similarity 30.8%; Pred. No. 3.3e-104;
Matches 277; Conservative 162; Mismatches 358; Indels 102; Gaps 23;

QY 16 SCCW---AFACHSTRESSPD--FTLPDYLACGLPPLHSGCL---QVRRHREYVTLCDRSC 66
DB 5 SCCWLLALWHTSAVCPDPAOKKGGIILGGLPHFGVAAKODIKSRESVECTR-- 62

QY 67 SFNEGHYLFQAMRIGVEYEINNSTALLPNTTLGYQLDVCSDSANVYATLRLSPGOHH 126
DB 63 -YNRGRFMQAMIFALEEINSSPALLPNTLIGYRIFDTCTVSK--ALEATLSPVANO 119

QY 127 IEIAG-----DLHYSTPVLAVIGPDSTNRAATTAALISPLVLMISTYASSELVSKRQ 181
DB 120 IDSINLDFPCNCEHIPSTIAVVGATGSGVSTAVANLGLFYIPVSVASSRLLSNKQ 179

QY 182 YPSFLRTIPNDKYOVETVLLQKRGWTWISLVGSSDDYDYGOLALPNOATGOSICIAF 241
DB 180 FKSFLRTIPNDHQATADITETFRMNMVGTIAADDYGRGLKEFEEDBERDICI 239

QY 242 KDIMPFSQVODERNOCIMRLHAOGATVAVVVFSSRQLARFEFESVLTNLTKWVASE 301
DB 240 SELT--SQYSEEEIQHVEVIONSTAKIYIVFSSGPLEPLIKETVARNITGKWLASE 297

QY 302 AMALSRHTGVPGIQRIGMVLGVAIQKRAVPLAF-----EAVYRADKEADRP 351
DB 298 AMASSSLAMPQYFHVVGITGFALKAQOIPGFRFLKVPKRSVHNGFKEFEETFN 357

QY 352 CH-----KG-----SMSSNQLCEQAOFAHMTPLKAFMSMSA 386
DB 358 CHLOGAKPLPVDTFLRGHESGDRSNSSTARPLCTGDNISVTPYIDTDLHLS 417

QY 387 YNAYRAVYAVAHGLHQLGC-----ASGACSR-GRVYPMQLLEQIHKVHFLHK-DTV 437
DB 418 YNVLAVYSIAHALQDIYTCLPGRGLFTNGSCADIKKVEAMQVLEKHLHMLPTNMGEQV 477

QY 438 AFNDNRDPLSSYNTIAMDNGPKMTFTVYLGSTMSPV-----QMINETKIQMH 486
DB 478 TFDECGDLVGNYSIIMNHLSPED-----GSIVREKGVYVYAKKGRFLINEKILMS 531

QY 487 GKDN-----QVPRKVSQDLEGHOR-VYTGFFHCOECVPCGAGTFUNKSDLYR 535
DB 532 GFSEPLFLVSLVQVPSNCSRDCLACTRKGLIEGEPTCCFEVECDGEDYSDTIDSA 591

QY 536 COPCKEEMAPESQOTCPRTVFLALREHTSWVLANTLLLLTAGLFAWHLDT 595
DB 592 CNKCPDDEMSNENHTSIAKEIEFLSWTEPRGIALTLFAVIGIFLAVLGVIFRMT 651

QY 596 VVRAGGGLCLMGSLAAGSGSLXGFGEPTRRACILRQALFALGFIIFLSCLTVRSQ 655
DB 652 IVKATNRLSYLLFSLCPSRSSLFFIGEPQDWTCLRQAPFISFVLCTISCLIVKTR 711

QY 656 LIIFKFEKTYPTFYH-AMVONHGAGLFVMISSAQAOLICTLVMPPLAREQRP 714
DB 712 VLWFE-AKIPTSHRKMWGLNLOFLVFLCTFMQIYICVWLTAPSSRNLEDE 769

QY 715 LVMECTETN--SLGFIAPLYNGLISAPACSYLGDLPENYNEAKCVFSLLEFVS 772
DB 770 IIFTCHEGSLMALGFLIG--YTCLLAICFFFAKRSKRLPENNEAKFITFSMLIF 827

QY 773 WIAFFT-ASYVDKYLPAAMMAGLSLSSGFGYFLPKCYVILCRDLNSTEHPQS 830
DB 828 WISFIPAYASTY-GKFSVAVEVIALILASFGILACIFPNKIYIILFKPSRMTIEVRCS 885

RESULT 10
US-08-480-751-6
Sequence 6, Application US/08480751
Patent No. 5858684
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:


```

; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 209/069
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1088 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-353-784-6

Query Match      25.2%; Score 1130; DB 3; Length 1088;
Best Local Similarity 30.8%; Pred. No. 3.3e-104;
Matches 277; Conservative 162; Mismatches 358; Indels 102; Gaps 23;

OY 16 SCCW---ARACHTESPDD--FLPBDYLLAGFLPHLSGL-----QVRHREPTVLCDRSC 66
    ||||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 5 SCCWVLLATWHTSAVGPQRAQKGDILLGLFPIHFGVAARDODLKSRPESVEGIR-- 62
    ||||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 67 SENEGYHLFOAMRLGVEEINNSTALLPNTLIGLYOLYDCSDSANYATLRVLSPGOHH 126
    : : : : : || : : : : : ||||| : : : : : || : : : : : || : : : : : || : : : : :
DB 63 -YNGRFRLQAMIFAEIENSSPALLPNTLCYRIFDTCNTYK--ALEATLSFAQNK 119
    : : : : : || : : : : : ||||| : : : : : || : : : : : || : : : : : || : : : : :
OY 127 IELQ-----DLHYSPTVAVIGPDTNRAATTAALLSPFLVPMISYAASETLSVKRK 181
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 120 ISLNLDFRCNCEHPISTLAVVAGATGSGVSTAVANLLGFYIPOVSYASSRLSKNKQ 179
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 182 YPSFLTPINDKQOVEMVLLQKPGMTWISLVGSSDDDGQGLQLENQATOGGCIAT 241
    : : : : : ||||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 180 FNSFLTPINDKQATAMADILEYFRMNMVGTIAADDDGPGIEFRFEAEERDIDF 239
    : : : : : ||||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 242 KDIMPFAQVGERMOCMLNHLQAAGATVVVFSRQLARVFESVYLTMIGKVVASE 301
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 240 SELI--SQSDEEIKHVEVIONSTAKYLVVSSGPDLEPLKEIYRNITGKIMASE 297
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 302 ANAASHITGVGICIGIWLGVATOKRAVPGIKAF-----EAVARADKEAPRP 351
    || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 298 AMASSSLIAMPOYFHVVGIGIFALKAGQIPGFRFLKVKHPRKSYHNGKAFWEETEN 357
    || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 352 CH-----KG-----SMCSSNOLCRCOAPMAHTMFKLAFASSSA 396
    || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 358 CHIOBAKGPLPVDTFLRGHEESGDRFSNSTAFRPLCTGDENISSEVETPYIDYTHLRIS 417
    || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 387 YNAARAVYAVAHGLHQLAG-----ASGACSR-GRVYPMOLLEQIHVHFLHK-DIV 437
    || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 418 YNYVLAVYSIAHALQDIYICLPGRLFTNGSCADIKKVEAMQYKHLRHLPNNMGQV 477
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 438 AFNDNDPLSSYIIAMDNKGPMTFTVLGSSSTWSPY-----QINNETKIQH 486
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 478 TDECDGLVNGSIIIMHLSPED-----GSIYFKEGYNYVYAKKGERLFINKEKIMS 531
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 487 GRDN-----QYKSVSSDCLEGHOR-VVTGFHHCCEFCVPCGAGTFLNKSDLYR 535
    || : : : : : ||||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 532 GPSREPLTVLSVLYQVPSMCSRDCLAGIRKGIIEGEPCCFECEVCPDEYDETDASA 591
    || : : : : : ||||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 536 COPCGKEEAPESQTCFPTTVFLALREHTSVWLLAANTLLLLLTAGLAFAMHLDP 595
    || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 592 CNKCPDPMFSENHTCIKKEIFELSWTEPFGIALTLFALVGLFALFVGLFIRPNP 651
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 596 VYRAGRGLCFMLGSLAAGSSLYGFPGPTPAQLRQALPAGLTITLSCLTVRSFQ 655
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 652 YKATNRELSTYLLFSLCCFSSSLFEIGEPQDWTCRLRQPARGISFVLICISILVKTNR 711
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 656 LIIFKSTKVPFTFYH-AWYQNHGAGLFWMISSAOLLIIMLVWYTPLPAREYQFPH 714
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 712 VLVVEE-AKIPISHRKMGMLNLOFLVFLCTFMQIVICVIMLYTAPSSSYRNOLEDE 769
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :

```

```

OY 715 LVMLECTETN--SLGFILAFLYNGLLSISAFACSYLCKDLPENYNEAKCVTFSLLENFVS 772
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 770 IIFITCHESSLMALGFILG--YTCILAIICFFAFKRSKRKLPENFNEAKFTTFPSMLIFFIV 827
    : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
OY 773 WIAFFT--ASVYDGKYLPAANMAGSLSSGGFGYFLPCVYILCRPDINSTPHRQAS 830
    || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
DB 828 WISFIPAVASTY-GKFEVSAVEVIAIILASFGLLACIFPNKITYIILLFKPSRNTIEEVACS 885
    || : : : : : || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :

RESULT 13
US-08-484-719B-6
; Sequence 6, Application US/08484719B
; Patent No. 6031003
; GENERAL INFORMATION:
; APPLICANT: Edward F. Nemeth, Edward M.
; APPLICANT: Brown, Steven C. Hebert,
; APPLICANT: Bradford C. Van Wageningen,
; APPLICANT: Manuel F. Balandrin,
; APPLICANT: Forrest H. Fuller, Eric G.
; APPLICANT: Delmar, Scott T. Moe
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS Word
; SOFTWARE: FastSeq for Windows Version 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,719B
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Douglas C. Murdoch
; REGISTRATION NUMBER: 37,549
; REFERENCE/DOCKET NUMBER: 213/007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1088 amino acids
; TYPE: amino acid

```



```

QY 67 SENEHGYHLEOAMRLGVEINNSTALPNTITLGYOLYDVCSDSANVATLRLSLPGOH 126
DB 63 -YNFGFRLQAMTFAIEEINSSPALPLRLTGLRIFEDCNVSK--ALEATLSYFAQMK 119
QY 127 IELQ-----DLHYSPTVLAIVGPDSTNRATTAALLSPFLVPMISVASSSETLSVKRO 181
DB 120 IDSLNLMDEFCNCSSEHIPSIAVAVGATGSGVSTAVANLLGLFYIPQVYASSSSRLSNKNQ 179
QY 182 YPSFLRTIPNDKYOVETWLLLOKFGMTWISLVGSSDDYGOGLVQALEMOATGOGICAF 241
DB 180 FESFLRTIPNDHQATAMADITEFRMNVGTIAADDDGRPGIEFRFEAEERDICIPE 239
QY 242 KDIMEFSAQVGERMOCLMRHLAQAATVYVSSRQLARVEFESVLTNLTKWVASE 301
DB 240 SELI--SQSDEEIOHVEVIONSTAKIYVVFSSGPDLEPLIKEIVRNITGKILASE 297
QY 302 AMALSRHTGVGIGRIGVNLGVAIOKRAVPLKAF-----EEAYARADKAPRP 351
DB 298 AMASSSLAMPQYFHVVGGTIGFALKAGIIPGREFLKVHBRKSVHNGFAKEFEWETEN 357
QY 352 CH-----KQ-----SMCSNOLCRECOAFMAHTMPKLFKAFMSA 386
DB 358 CHLOGAGAPLVDYFRLGRHESSGDRFSNSTAFRLCTGDNISSEVETPYIDYHLRIS 417
QY 387 YNAYRAVAVAHGLHOLLG-----ASGACSR-GRVYPMOLLEQIHKVHFLHK-DTV 437
DB 418 YNAYLAIVYSAHALODIYCLPGRGIFTNGSCADIKKIVAMOVILKRLHNLFTNNMGQV 477
QY 438 AFNDNRDPLSSNYIIAMDMNGPKMTFTVIGSSTSVPY-----QININETIOH 486
DB 478 TFDECGDLVGNSTINMHLSPED-----GSTVFEVGYVNYAKKGERLFTNEEKILMS 531
QY 487 GKN-----QVPKSVCSDELEGHOR-VVTGFHHCCECYVPCGAGTFLNKSIDLX 535
DB 532 GSRBRLPFLVSLQVFPSCNRDCLAGRKIIIESEPCECEVCBPCGEVSDERDASA 591
QY 536 COPCKEMAPRGSOCTCFPRIVFPLAREHTSMVLIANTLTLILLTAGLFAHMLDTP 595
DB 592 CNKCDDEFSNENHSCIAKEIEFLSWTEPFGIALTLFVNLGIFLAFLVAGVFEIKRPNP 651
QY 596 VYRSAGRCFLMGLSIAAGSIVGFGEPRPRACLLNOALFALGFTIFLSCTVRSQ 655
DB 652 YKATNRELSTYLLSLCCFSSSLFGEPODWTCRLQRPAGISFVLCISLITAKTKR 711
QY 656 LIIIFKSTKVPFYH-AWVQNHGAGLFYMISSAQLICTLWLVWTFPLPAREYORFPH 714
DB 712 VLVFE--AKIPTSPHRKMWGLNLQFLVFLCTPMQIVICVIMLYTAPSSVRNDELE 769
QY 715 LVMECTEN--SLGFILAFLYNGLISAFACSTYLGKDPENYNEAKCVTFSLEFVYS 772
DB 770 IIFICHESLMAFLIG--YTCLLAACFPFAKSRKLPENFNFAKIFTESMLFFIV 827
QY 773 WIAFFT-ASYVDKYLPANMAGISLSGFGVFLPKCVIIICRPLNSTEHQAS 830
DB 828 WISFLPANTASTY-GKFSVAVEVIALIASFGLLACIFENKIYIILFKPSRNTIEVRS 885

```

RESULT 15

US-09-134-513-2

Sequence 2, Application US/09134513

Patent No. 6210964

GENERAL INFORMATION:

APPLICANT: Brown, Edward M.

APPLICANT: Diaz, Ruben

APPLICANT: Bai, Mei

APPLICANT: Quinn, Stephen J.

TITLE OF INVENTION: The Avian Extracellular Calcium-Sensing

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESS: Vinson & Elkins L.L.P.

STREET: 1455 Pennsylvania Avenue, N.W.

```

CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,513
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: BR1331/13003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)639-6604
TELEFAX: (202)639-6585
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-09-134-513-2

Query Match 25.1%; Score 1127.5; DB 4; Length 1059;
Best Local Similarity 30.7%; Pred. No. 5,6e-104;
Matches 276; Conservative 163; Mismatches 352; Indels 107; Gaps 24;

QY 14 LISC-----WAFCHSTESSPDITLPGDYLLALFLHSCLE-----QVRHREPVTL 62
DB 3 LYSCLILLFTWNTAAYGPNOQAQ--KKGDIIILGGLFPIHGVAAKDDLSRPSVBC 60
QY 63 DSCSSENEHGYHLEOAMRLGVEINNSTALPNTITLGYOLYDVCSDSANVATLRLSLP 122
DB 61 IR---YNFGFRLQAMTFAIEEINSSPALPLRLTGLRIFEDCNVSK--ALEATLSY 115
QY 123 GOHIELOQ-----DLHYSPTVLAIVGPDSTNRATTAALLSPFLVPMISVASSSETLS 177
DB 116 AONKIDSLNLMDEFCNCSSEHIPSIAVAVGATGSGVSTAVANLLGLFYIPQVYASS 175
QY 178 VKROYSPFLRTIPNDKYOVETWLLLOKFGMTWISLVGSSDDYGOGLVQALEMOATGOGI 237
DB 176 NKNQFSPFLRTIPNDHQATAMADITEFRMNVGTIAADDDGRPGIEFRFEAEERDI 235
QY 238 CIAFKDIMEFSAQVGERMOCLMRHLAQAATVYVSSRQLARVEFESVLTNLTKWV 297
DB 236 CIDFSELI--SQSDEEIOHVEVIONSTAKIYVVFSSGPDLEPLIKEIVRNITGKIL 293
QY 298 VASEANALSRHTGVGIGRIGVNLGVAIOKRAVPLKAF-----EEAYARADK 347
DB 294 LASEMASSSLAMPQYFHVVGGTIGFALKAGIIPGREFLKVHBRKSVHNGFAKEFEW 353
QY 348 APRPCH-----KGSNCSSN-----QLCRCQAFAHTMPKLFKAFMSA 383
DB 354 EFNCTLPSESKNSPASASPFHKAHEBGLAGNGTAARPPCTGDNITSEVETPYMDFTL 413
QY 384 SSAYNAYRAVAVAHGLHOLLG-----ASGACSR-GRVYPMOLLEQIHKVHFLHK- 434
DB 414 RISTYNYLAIVYSAHALODIYCLPGRGIFTNGSCADIKKIVAMOVILKRLHNLFTSN 473
QY 435 DTVAFNDNRDPLSSNYIIAMDMNGPKMTFTVIGSSTSVPY-----QININETRI 483
DB 474 EYVDFDEFGDLVGNSTINMHLSPED-----GSVYFEVGYVNYAKKGERLFIENKTI 527
QY 484 QNHGKNQVPSKVSDDCLEGHOR-VVTGFHHCCECYVPCGAGTFLNKSIDLXROPCGKE 542

```


PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;

PI Zambrowicz B, Sands AT;
 XX MPI: 2001-616474/71.
 DR N-PSDB: AAD19501.
 XX
 PT Novel isolated polynucleotides encoding human G protein coupled
 receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 expressed NGPCRs for diagnosis of disease, and as probes or primers
 XX
 PS
 Claim 3: Page 72-74; 80pp; English.

The present sequence is human novel G-protein coupled receptor
 (NGPCR) protein. NGPCRs are transmembrane proteins that span
 the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heart rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.

XX
 XX Sequence 841 AA:

Query Match 100.0%; Score 4493; DB 22; Length 841;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLCTARLVGLQLLISCCMAFACHSTESSPDTLPDGLYLAGLPLHSGCLQVRRPEVT 60
 DB 1 mlctarlvglqlilisccwafachstesspdtlpgdyllaglfphsgclqyrhpervt 60
 QY 61 LCDRSCSPNKHGHLFOAMBLGVEETINNTALLPNTIGQGLDVCSDSNVYATLRVLS 120
 DB 61 lctdrscspnkhghlfoamblgveetinnstallpntllygqldvcsdsanvyatlrvis 120
 QY 121 LPSGHHIEIQLDGLAHSPYLAIVIGPDSNRATTAALSPFIYPMISYAASETLSVKR 180
 DB 121 lpsghhieqlldglahspylavigpdsnraatlaalspflypmisyaasetlsvkr 180
 QY 181 QYPSLRTIPNDKYQVEYVWLLIQKFGWTWISLVSSDDYGLGVQALENQATGGGICIA 240
 DB 181 qypslrtipndkyqveywlllqkfgwtwislvsddyglgvqalemqatgggicla 240
 QY 241 FKQIMFSAQVGERMOCIMRHIAOGATVYVVFSSROLARFEESVLTNLTKGWVMS 300
 DB 241 fkdimpfsaqvgermqcmrthlaqagatvvvfssrqrlarfeesvltnltkgwvms 300
 QY 301 EAWALSRRHTTGVGIORIGVNLGVAIOKRAVPGIKAFEEAYARADREAPRCHKSGWSS 360
 DB 301 eawalsrrhttgvgiorigvnlgvaiokraavpgikafeeayaradrepchkgswss 360
 QY 361 NOLCRECOAFMAHTMPKLAFSMSAYNAVRAVVAHGLHQLLGASGASGRVYPMW 420
 DB 361 nqlcrecoafmahtmpklafsmsaynavravvahnghlqlgscasgcsgrvympw 420
 QY 421 LLEQIHKVFILHKDVAFNDRDPLSSYNIITAMDNGKRWTFVYVYSSYVQNLIN 480
 DB 421 lleqihkvfihkdvafndrdplssyniitamdngkrwtfvyvyssyvqnl 480
 QY 481 TKIQHGNKNOVPRKSVCSDDLGEHQVYVGFHHCCECPGAGAFGLNKSDDYRCPG 540
 DB 481 tkiqhgnknovprksvcsddclgehqvvyvghhccfecpccgagfllnksdlyrcpg 540

QY 541 KEEMAPESGQCEPRTVVEFALREHTSWLLAANTLLLLLTAGLAFAMHLDTPVYRSA 600
 DB 541 keemapesgqceprtvefalrehtswllaanllllltaglfamhldtpvyrsa 600
 QY 601 GGRLCFLMGLSLAAGSGSLXGFFGEPTRRACLLRQALFALGFTIRFSLCTVASFQIITF 660
 DB 601 ggrlcfmlgslaaagslxfgeptrracllrqalfalgtirfslctvassfqiitf 660
 QY 661 KFSTKVPTFYHAWVONHGAFLFYMISAAQLICLTWLVWMPPLPAREQRPFLVMEBC 720
 DB 661 kfstkvpftyhawnhgaflfymisaaqlcltwlvwmpplpareqrpfhvmlebc 720
 QY 721 TETNSLGFILAFLYNGLSLISAFACSYLGKDLPEYNNEAKCVTFSLFNFVSMIAFETTA 780
 DB 721 tetnslgfilafllynslisafacsylgkdlpeynneakcvtfslfnfvsmiafetta 780
 QY 781 SVYDGKYLPAANMAGLSLSSGFGYFLPKCVILCRDNLNSTERFQSIQDYTRFGSS 840
 DB 781 svydgkylpaanmaglsissgfyflpkcvylcrpdlnstehfgasldqytrfcs 840
 QY 841 T 841
 DB 841 t 841

RESULT 2
 AAE10372
 ID AAE10372 standard; Protein; 841 AA.
 XX
 AC AAE10372;
 XX
 DT 10-DEC-2001 (first entry)
 XX
 DE Human taste receptor, hTIR1 protein.
 XX
 KW Human; taste-cell-specific G protein-coupled receptor; hTIR1; drug;
 KW genetic modulation; pharmaceutical; taste sensation; food industry;
 KW chemosensory transduction.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 274 /note="Encoded by TTTT"
 FT
 PN WO200166563-A2.
 XX
 PD 13-SEP-2001.
 XX
 XX 07-MAR-2001; 2001WO-US07265.
 PF
 XX 07-MAR-2000; 2000US-0187546.
 PR 07-APR-2000; 2000US-0195336.
 PR 06-JUN-2000; 2000US-0209840.
 PR 23-JUN-2000; 2000US-0214213.
 PR 17-AUG-2000; 2000US-0226448.
 PR 03-JAN-2001; 2001US-0259227.
 XX
 PA (SENO-) SENOMYX INC.
 XX
 PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
 XX
 DR MPI: 2001-582267/65.
 DR N-PSDB: AAD17515, AAD17516.
 XX
 PT New mammalian taste-cell-specific G protein-coupled receptor
 PT polypeptides for identifying compounds that modulate taste signalling
 PT are useful in food, to modulate the sweet taste of foods or drugs
 XX
 PS Claim 152; Page 83; 119pp; English.
 CC The invention relates to mammalian taste-cell-specific G protein-coupled
 CC receptors, TIR and their corresponding cDNA molecules. Taste receptors,

CC T1R are useful for screening compounds which are used to activate or
 CC modulate chemosensory transduction, such as taste sensation. The
 CC identification and isolation of novel taste receptors and taste
 CC signalling molecules allow for new methods of chemical and genetic
 CC modulation of taste transduction pathways. The taste modulating
 CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. T1R,s are also useful in
 CC biochemical assay for identifying tastant (T1R) ligands having binding
 CC specificity for T1R involved in taste signalling. The present sequence is
 CC human taste-cell-specific G protein-coupled receptor, hT1R1 protein.
 XX
 SQ Sequence 841 AA;

Query Match 99.98; Score 4489; DB 22; Length 841;
 Best Local Similarity 99.98; Pred. No. 0;
 Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLCCTARLVGIQLITSCCWAFACHSTESSPFTLPGLVLAAGLPPLHSGCLQVRHREPT 60
 DB 1 mlctarlvgllqlliscwafachstesspftlpdgyllagllfpblhsgclqvrhrpevt 60
 QY 61 LCDRSSCFNEHGYHLFQAMRLGVEEINNSTALLPNITLGQLYDVCSDSANVATLRVLS 120
 DB 61 lcdtscsfnehgylhfqamrlgveeinntallpnitlgylqlydvcsdsanvyatlrvls 120
 QY 121 LRGQHITFLOGDLAHSTYVAIVGPDSTNRATTAALSPFLVPMISAASEFLSVKR 180
 DB 121 lrgqhiteflogdlahtstvyavivgpdstnraatlaalspflvpmisyaasaelsvkr 180
 QY 181 QYPSFLRTIPMDKYOVEMVLLLOKFGWTWISLVGSSDDYDQLOLQVQALQENATGGGICIA 240
 DB 181 qypsfllrtipmdkyovetmvllokfgywtwislvgssddydyqalqenatqggicia 240
 QY 241 FKDIIMPESAQYDDEMOCLMRHLAOGATVVVVSRLARFESVVLNLTGKRWVAS 300
 DB 241 fkdimpesaqydgdermoqlmrhlaogatvvvvsrqlarfesvvlntltgkvwvas 300
 QY 301 EAMALSRITITVPGIORIGMYLGAIVQKRAVPGIKAFEEATARADKEAPRCHKSGWSS 360
 DB 301 eamalsrilitvpgiorigmylgaivqkravpgikafeeatarakkaprchksgwss 360
 QY 361 NQLRECOAFMAHMPKLFKAFSSSAYNAVYAVAHAGLQLLGCAGASRGVYVWMO 420
 DB 361 nqlrecoafmahmpkfkafsssaynavyavaahagllqllgcaagsrgvyvwm 420
 QY 421 LLEQIHKVFLLHKDTVAFNDNRDPLSSYNIADWMNGPKMTFTVLGSSSTSPVOLNINE 480
 DB 421 lleqihkvfllhkdtvafndnrplssyniadwmngpkmtftvlgssstspvolnine 480
 QY 481 TKIWMGKDNQOVKSVSSDCLBQHORYVTFHHCCFECVPCGAGTFLNKSDLYRCOPCG 540
 DB 481 tkiwmgkdnqovksvssdclbqhoryvtfhhccfecvpcgagtfllnkstdlyrcopcg 540
 QY 541 KEKAPAGSQCPRFTVETALREHTSMVLLAANTLILLILLAGAGFAMLLDPVYVSA 600
 DB 541 kekapagsqcpftvetalrehtsmvlllaantlillililagagfamlldpvvysa 600
 QY 601 GGRICFTLMLAGSGLSYGFEEPTRPACILRQALFALFTTFLSLCTVRSFOLLITF 660
 DB 601 ggricftlmlagsglsygfefptrpacilrqalfalfttflslctvrsfollitf 660
 QY 661 KESTRVPTFYHAWVQNHGAGLFVMISSAQLILICTLWLVVTPPLPAREYQRPFLVMEC 720
 DB 661 kestrvtptyhawvqnhgaglfvmisaaqlilictlwlvvtplpareyqrfpvlmlec 720
 QY 721 TETNSLIGTILAFYNGILSTISAFACSVYGRDLPENYNAKVTSTSLFENFSVATFTTA 780
 DB 721 tetnsligtilyafyngilstisafacsvygrdlpennyakvtsstslfenfsvatftta 780
 QY 781 SVYDGKYLPAANMAGLSLSSGFGYFLPKCYVILCRPDNLSTEHFQASIQDTRRCGS 840
 DB 781 svydgkylpaanmaglslessgfgyflpkcyvilcrpdnlstehfqasidqdyrrcgs 840

DB 781 svydgkylpaanmaglslessgfgyflpkcyvilcrpdnlstehfqasidqdyrrcgs 840
 QY 841 T 841
 DB 841 t 841

RESULT 3
 ID AAY45023 standard; Protein: 777 AA.
 AC AAY45023;
 DT 31-MAY-2000 (first entry)
 DE Human sensory transduction G-protein coupled receptor-B3.
 KW Human; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX FT Misc-difference 289
 XX FT /note= "Encoded by AC"
 XX PN W0200006592-A1.
 XX PD 10-FEB-2000.
 XX PE 27-JUL-1999; 99WO-US17099.
 XX PR 28-JUL-1998; 98US-0094465.
 XX PA (REGC) UNIV CALIFORNIA.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;
 DR WPI: 2000-205451/18.
 DR N-PSDB: AAZ50745.
 XX PT New isolated sensory transduction G-protein coupled receptor, useful
 XX PT for developing products for use in studying and modulating the taste
 XX PT transduction pathway -
 XX PS Claim 22; Page 76; 83pp; English.
 XX CC The present sequence is a taste cell specific G-protein
 CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
 CC This sequence was isolated from the human testis library.
 CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
 CC lower expression in circumvallate taste receptor cells of the tongue.
 CC The present sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 XX
 SQ Sequence 777 AA;

Query Match 90.68; Score 4068.5; DB 21; Length 777;
 Best Local Similarity 96.68; Pred. No. 0;
 Matches 767; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 64 RSCSFNEHGYHLFQAMRLGVEEINNSTALLPNITLGQLYDVCSDSANVATLRVLSLPG 123
 DB 64 rscsfnehgylhfqamrlgveeinntallpnitlgylqlydvcsdsanvyatlrvlsldpg 123

```

Db      1 rscsfnehyhlfqamrlgveelinstallpnltyglydyvcsdsanyatlrvlslp 60
Oy      124 OHHIELOGDLHYSPVLAVIGDPDSTNRATTAALISPLVPMISYAASESTLSVROFP 183
Db      61 qnhleqollhyspvcavlgpdsctnaactaallspflv-hlisyaaasetlsvrqp 119
Oy      184 SFLRTIPNDKYOVEWVLLQKFGWTWISLVGSSDDYGQLGVQALENQATGOGICIAFD 243
Db      120 sflrtipndkyvelmvlallqkfgwtwislvgssddyqglvgalenglvvgiclaefd 179
Oy      244 IMPFSQVQVDEWOCIMRLAQAATVVVVFSSROLARFFESVLTNTLGKVAASEM 303
Db      180 lmpfsaqvqdermqclmrllaqagatvvvffsrlarvfesvltntlgkvwaseaw 239
Oy      304 ALSRHTGVPGIORIGMVGVAIOKRAVPGLKAFEEAYARADKEARPCPKSGWSSNOL 363
Db      240 alsrhtgvpgiorigmvagvaqkravpjlkafeeaaradkearpcpksgwssnql 299
Oy      364 CRECOAFMAHTMPKTKAFSSSAYNAYRAVAVAHGLHOLLCCASGACSRGRVPMQLE 423
Db      300 crecgaftmahmpkltkafssasaynayravahglhllgcaselcgrvypwqlle 359
Oy      424 QIHKVFELHAKDPVANRNDRLPLSSNITIAMDMNGPKMTFTVLGSSSTSPVOLNINET 483
Db      360 qihkvflllhkdtvalndrplssynllawdmngpkmtftvlgssstspvolninetk 419
Oy      484 QMHGKDNQPKSVCSDDCEGHQRYVTGFHHCCEFCVPCGAGTFELKNSDLRYQPCGKEE 543
Db      420 qmhgkdnqpksvcsddceghqrvvtgfhccfecvpcgagtfelknselyrcpgqtee 479
Oy      544 WAPESQTCFPTVFLALREHTSWVLAAATLLLLLTAGLAFWMLDTPVVSAGR 603
Db      480 wapesqtcfptvflalrehtswvllaantllllltagllfawhldtpvvsagr 539
Oy      604 LCFIMGSLAASGSLYGFEGETRPACLLROALFALGTIFLSCLTSSFOILIEKES 663
Db      540 lcfimgslaasgslygfegetrpacllrqalialgtiflsccltssvfglllffks 599
Oy      664 TRVPTFFYHAWONHGALFVMTSSAQLLICTLWLVWTPLPAREYQRPPLMLECTET 723
Db      600 tlvptfthawvqnhgagllvmtssaaqlllcltlwlvwtplpareyqrpplmlectet 659
Oy      724 NSLGFIAFLVNGLSISAPACSYLGKDLPEYNNEAKCYTFSILNFVSMIAFTTASVY 783
Db      660 nslgfiaflvnglslsapacsylgkdlpeynneakcvtfslfnfvswiafftasy 719
Oy      784 DCKVLPAAAMMAGLSLSSFGGYFLPKCYVILCRPDLNSTEHFOASIDYPRCGST 841
Db      720 dckvlpaaammaglsissfggyflpkcyvllcrpdlnstehfgasidgytrcgst 777

RESULT      4
AAEL1970
ID      AAEL1970 standard: Protein; 763 AA.
XX
AC      AAEL1970;
XX
DT      18-DEC-2001 (first entry)
XX
DE      Human novel G-protein coupled receptor (NGPCR) protein #2.
XX
KW      Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW      antisense-therapy; signal transduction; behavioral disorder; obesity;
KW      heartbeat rate; inflammation; immune disorder; diabetes; cancer;
KW      coronary disease.
XX
OS      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      Misc-difference 29 /note= "Encoded by TTT"
FT      Misc-difference 294 /note= "Encoded by RCA"
FT

```

```

XX      XX
PN      WO200172842-A2.
XX      XX
PD      04-OCT-2001.
XX      XX
PF      28-MAR-2001; 2001WO-US09996.
XX      XX
PR      28-MAR-2000; 2000US-192978P.
XX      XX
PA      (LEXI-) LEXICON GENETICS INC.
XX      PI      Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
XX      PI      Zambrowicz B, Sands AT;
XX      PI      WPI: 2001-616474/71.
XX      PI      N-PSDB; AAD19502.
XX      PS      Novel isolated polynucleotides encoding human G protein coupled
XX      PT      receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
XX      PT      expressed NGPCRs for diagnosis of disease, and as probes or primers
XX      PS      Claim 4: Page 75-76; 80pp; English.
XX      CC      The present sequence is human novel G-protein coupled receptor
XX      CC      (NGPCR) protein. NGPCRs are transmembrane proteins that span
XX      CC      the cellular membrane and are involved in signal transduction after
XX      CC      ligand binding. The NGPCR polynucleotide sequences, are useful
XX      CC      in diagnosis and treatment of a disease involving NGPCR, for detecting
XX      CC      mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
XX      CC      disease, for screening drugs effective in treatment of symptomatic or
XX      CC      phenotypic manifestations of perturbing the normal function of NGPCR in
XX      CC      the body or abnormalities in the signal transduction pathway mediated by
XX      CC      NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
XX      CC      trial monitoring and/or the treatment of physiological (heartbeat rate)
XX      CC      or behavioural disorders. NGPCR is useful for identifying compounds
XX      CC      useful in the therapeutic treatment of obesity, inflammation, immune
XX      CC      disorders, diabetes, heart and coronary disease, metabolic disorders and
XX      CC      cancer. Labelled NGPCR nucleotide probes can be used to screen a human
XX      CC      genomic library which is helpful for identifying polymorphisms,
XX      CC      determining the genomic structure of a given locus/allele and designing
XX      CC      diagnostic tests.
XX      SQ      Sequence 763 AA:
XX
Query Match      90.3%; Score 4058; DB 22; Length 763;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 763; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      79 MRLGYEINNSYALLPNTITLGYQLYDVCSDSANYATLRLVLSLPGOHIELOGDLHYSP 138
Db      1 mrlgyeelnstallpnltyglydyvcsdsanyatlrvlslpghlelsgdillhysp 60
Oy      139 TYLAVIAGDPSTNRATTAALISPLVPMISYAASESTLSVROFPSELTINDKYQVET 198
Db      61 tylvavipdstnraactaallspflvpmisyaasetlsvkrqypsflrtipndkyqvet 120
Oy      199 NVLLQKFGMTWISLVGSSDDYGQLGVQALENQATGOGICIAFKDIMPFSQAQVDERMQC 258
Db      121 nvllqkfgwtwislvgssddyqglvgalengatvggiclaefkdtimpfsaqvdermqc 180
Oy      259 LMRHLAQAATVVVVFSSROLARFFESVLTNTLGKVAASEMALSHTITGVPGIORI 318
Db      181 lmrhlaqagatvvvffsrlarvfesvltntlgkvwaseawalshtitgvygigr 240
Oy      319 GNLGVAIQKRAVPGLKAFEEAYARADKEARPCPKSGWSSNOLCRECOAFMAHTMPK 378
Db      241 gnlgyaigkravpjlkafeeaaradkearpcpksgwssnqlcrecgaftmahmpk 300
Oy      379 KAFSSSAYNAYRAVAVAHGLHOLLCCASGACSRGRVPMQLEQIHKVFELHAKDPVA 438
Db      301 kafssasaynayravavahglhllgcasgacsrgrvypwqlleqihkvfllhkdva 360

```


Oy	781	SVYQGYKYLPPANMAMGSSISGSGGGEFLPKRCVILLCRPLDINTEHFOASIDYPRGCS	84	00
Dd	780	SIYGSYSLPAAVNVIAGLTCLISGTSYFLPKCYVILLCPMLNTEHFGASIQYLTTCGL	839	
Oy	841	T 841		
Dd	840	t 840		
	RESULT	6		
ID	AAAY45026			
XX	AAAY45026	standard; Protein; 840 AA.		
AC	AAAY45026:			
XX				
DT	31-MAY-2000	(first entry)		
XX				
DE		Rat sensory transduction G-protein coupled receptor-B3 variant #1.		
XX				
KW		Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;		
KW		sensory cell; taste receptor cell; screen; taste modulator;		
KW		pharmaceutical; food industry; taste topographic map; tongue; variant		
XX				
OS	Rattus sp.			
OS	Synthetic.			
XX				
FT	Key	Location/Qualifiers		
FT	Misc-difference 33	/note= "Wild type Leu substituted by Ile"		
XX				
PN	WO200006592-A1.			
PD				
XX	10-FEB-2000.			
PF	27-JUL-1999;	99WO-0517099.		
XX				
PR	28-JUL-1998;	98US-0094465.		
XX				
PA	(REGC) UNIV CALIFORNIA.			
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.			
XX				
PI	Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;			
DR	WPI; 2000-205451/18.			
XX				
PT		New isolated sensory transduction G-protein coupled receptor, useful		
PT		for developing products for use in studying and modulating the taste		
XX		transduction pathway -		
PS				
XX	Disclosure; Page -; 83pp; English.			
CC				
CC		The present sequence is a polymorphic variant #1 of rat taste cell		
CC		specific G-protein coupled receptor, GPCR-B3 which is involved in		
CC		sensory transduction. GPCR-B3 is specifically expressed in foliate		
CC		and fungiform cells, with lower expression in circumvallate taste		
CC		receptor cells of the tongue. The protein has an extracellular domain		
CC		seven transmembrane domains and an intracellular domain.		
CC		The GPCR-B3 sequence is used to screen compounds that modulate sensory		
CC		signalling in taste cells, especially taste modulators useful in		
CC		pharmaceutical and food industries to customise taste. The sequence		
CC		can also be used as probe for identifying taste cells and		
CC		subsets of taste receptor cells such as foliate, fungiform and		
CC		circumvallate. Such probes are also useful to generate taste		
CC		topographic maps that elucidate the relationship between the taste		
CC		cells of the tongue and sensory neurons leading to taste centres		
CC		in the brain.		
CC		Note: The present sequence is not given in the specification but is		
CC		derived from rat GPCR-B3 sequence shown in page 75 (AAAY45021).		
XX				
50	Sequence	840 AA;		

Query Match	74.88;	Score 3359.5;	DB 21;	Length 840;
Best Local Similarity	74.08;	Pred. No. 0;		
Matches 622;	Conservative 82;	Mismatches 136;	Indels 1;	Gaps 1;

QY	1	MILCTARLVGKLLISCCMAFACSTSSPPFTLRGVLLAAGLRPLHSGLQVNRHREVT	60
Db	1	mlfwaahll-lslqlvycwaisscqrteespgfs:ipdgffllagflshlndgclqrthprlvt	59
QY	61	LCDRSCSFNEGHVHLFQAMRLGVEEINNSTALLPNTLTGQLDYDVCSDSANVTATLRVLS	120
Db	60	scdrpdsfngghyhlfgamrtfveeennsallpnltlgyelvdvcesanayactlrvla	119
QY	121	LPQGHIELGGDDLHYSPTVLAVIGPSTNKAATTAALLSPFLVPMVSYAASSETLSVKR	180
Db	120	lqgrhrlelqkdlrlnhskvvafliprpnthavtcaalllqrlfmlrvysaassvlsakt	179
QY	181	QYPSFLRTIPMDKYOVETMVLLLQKFGMTWISLVSGSSDDYGOGLQVALNQATGOGICIA	240
Db	180	kfrsflrtlvprddrhqvevmwqlqsfqvwvslsigsydgylqvgalelavrglcva	239
QY	241	FKDIPFEAGVGDERMOCMLNHLLAQAATVYVYSSNQALRAVFESVVLNLTGKVVVAS	300
Db	240	fkdlvrfsaayvdrpmqsmhglaqatlvtvvlnshrlavfrsvvlnlllqgkvvas	299
QY	301	EWALSRHITGVPGIQRIGMTLGAIGLQKRAVPGIKAEAEAYARADKRAPRCHKSGWSCS	360
Db	300	edwaistylstvtgdlqglvtlvaavqgrqyrgjkeeteesylvavtapaacregswast	359
QY	361	NOLRCSEAFMAHTWPKIKAFSMSGSAYNARAVYAAVAGHLQLLIGCAGSCSRGRYPMQ	420
Db	360	nqlcrechtlftrmptlrgafsmasaaayvyaavahglhqllygctseiscsrpyvpwq	419
QY	421	LLEQIHKKVHFLIHDQVAFNDPNDRLSSYNLIADMDMGKMTFVLGSSWSPVOLINE	480
Db	420	llqqlvkvnlhhenlvaiddngdclgyddllaawmgpewtlelvgasalspvhldlnk	479
QY	481	TKIQMGKDNQVPRKSVCSDDLCEGHORYVTGFIHNCCEFCVPCGAGTFLNKSDDLRCPCG	540
Db	480	tklqghgmngyprsvctcdclaghhvrvyghscfecpceagtlmmsehlhcpgs	539
QY	541	KEENABEGSQFPRPVVFLALREHTSMVLIAANTLLLLLLLTGAGLFAMNHLDPVYRSA	600
Db	540	teewarkestctcfprveflawherlsivliaantcllllllyvgtaglfawhfnhprvrsa	599
QY	601	GGRLCEFLMLGSLAAGSGSLYGFCEPFPAPACLEQALFALGFTFLFSLCIVRSFOLIIIF	660
Db	600	ggrrlcfllmlhgsilvsgscsfysfdeprtcpaclllqplfs:lgfaifslclstfrsfyvlif	659
QY	661	KFTSKVPPTFYAAWYONHAGLEFVMISSAODLLCLTLMVLVYTPRLPAREYORFPLVMLEC	720
Db	660	kfstkvprfftytwaqnbhagjfvilvstevhlhllcltwlvmtptrpreygrfplvlalec	719
QY	721	TETNSLGTLAFLVINGLISLAFACSYLIGKDLPENYNWADACVYFSLFLFNVSMTAFPTTA	780
Db	720	tevnsvagflaflthmlslstlscfcsylygkelpenyeakcvtslllnltvswtaiftma	779
QY	781	SVYDQKYVLPANMMAGLSSLSGSGFGVGLPFCYVILRCPDLNSTEHFQASIODYTRRCGS	840
Db	780	slvgsgsyipavnvlaaglttllsggfsyglfipkcyvllcrpelmntehgaaidsqdytlrrcgt	839
QY	841	T 841	
Db	840	t 840	
RESULT 7			
AAV45027 standard; Protein: 840 AA.			
AAV45027:			
31-MAY-2000 (first entry)			

```

DE Rat sensory transduction G-protein coupled receptor-B3 variant #2.
XX
KM Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
KW sensory cell; taste receptor cell; screen; taste modulator;
XX pharmaceutical; food industry; taste topographic map; tongue; variant.
XX
OS Rattus sp.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 84 /note= "Wild type Glu substituted by Asp"
FT WO20006592-A1.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-US17099.
XX
PR 28-JUL-1998; 98US-0094465.
XX
PA (REGC ) UNIV CALIFORNIA.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
DR WPI; 2000-205451/18.
XX
PT New isolated sensory transduction G-protein coupled receptor, useful
PT for developing products for use in studying and modulating the taste
PT transduction pathway -
XX
PS Disclosure; Page -: 83pp; English.
XX
CC The present sequence is a polymorphic variant #2 of rat taste cell
CC specific G-protein coupled receptor, GPCR-B3 which is involved in
CC sensory transduction. GPCR-B3 is specifically expressed in foliate
CC and fungiform cells, with lower expression in circumvallate taste
CC receptor cells of the tongue. The protein has an extracellular domain,
CC seven transmembrane domains and an intracellular domain.
CC The GPCR-B3 sequence is used to screen compounds that modulate sensory
CC signalling in taste cells, especially taste modulators useful in
CC pharmaceutical and food industries to customise taste. The sequence
CC can also be used as probe for identifying taste cells and
CC subsets of taste receptor cells such as foliate, fungiform and
CC circumvallate. Such probes are also useful to generate taste
CC topographic maps that elucidate the relationship between the taste
CC cells of the tongue and sensory neurons leading to taste centres
CC in the brain.
CC Note: The present sequence is not given in the specification but is
CC derived from rat GPCR-B3 sequence shown in page 75 (AAY45021).
XX
SQ Sequence 840 AA;

```

Db	180	kfgstlrvpsdrtnqveemwqllsgfswvwlslsgsytdgqgvgvgaaleelavprlctva	239
QY	241	FKDIMPSSAQGDERMQCLMRHNLQAQATVVVVVFSSQQLARVFESVYLTNLTKGYWVAS	300
Db	240	fkdlvprfaearvgdprmsmqmhlraqacttvtvvvfsmnhlrvffrsrvlanllgkywvas	299
QY	301	EAMLSRIHTGVPGIQRIGMWLGVAIQKRAVPGLKAREEAYAADKAPPCPKHSGMSCS	360
Db	300	edwastlytstvvgdggvgvlgvavvqrvqpglkefeesvaylavtaapsacpewscst	359
QY	361	NQLRCRCQAFNAHMPKIKAFSSMSAANAYRAVYVAVAHGHOLLGCAAGCAGSRGRVPMQ	420
Db	360	nqlrcrcchttftrmptlrgtmsaayrvvyeavavahghlqllgctselcsrcgryppmq	419
QY	421	LLEQIHAKHFLHLDYAFNDNRPLSSYNIANDMGRPMTTTLVGLSSMSRVOLINDE	480
Db	420	llqgylxvnhllhenlvafidngdltgyydliawdmwpewtellelgsaaispvhldlnk	479
QY	481	TKIQHGKDNQVPRKSSVSDCLEGHQRYVTGFIHHCCEFCVPCGAGFTLANKSDLYRCQPCG	540
Db	480	tkiqghgdnqgvprsvcttdclaghnrvvvgshhccefcvpcgaqgtlmselhdqpgv	539
QY	541	KEEWAPEGSQTCFPRYVFLALREHTSWLANTLLLLGLTAGLPAWHLDTPVYRGA	600
Db	540	teewapkestctcfprtveflamhepislvliaantlllllyvtaqlfawhfltpvyrsa	599
QY	601	GGRLCEFLMLGSLAAGSGKVGFECEPRPAPCLIKQALFALGFPIPLISCLVRSFOLITIF	660
Db	600	ggrlcelflmlgslavagsgsfysfdeprvpallllqplfsgfalflscldlrsfgylvllf	659
QY	661	KFTKVPTEPYAVAWYONHAGLEFVMISSAOLLICLTWLVVYTPLPAREYORFPHLWLEBC	720
Db	660	kftkvpfeytvaqwnbgagfivfsvstvlhllcltvlwmvtrpcreqtrfphylvleec	719
QY	721	TEINSGITLAEFLNGILISTAFACSVTLGKDLPENTVEACVPEFSLTFNFSVIAFEFTA	780
Db	720	teinsvglllafthmlhllstlrsfcsylgkelpenyneacvctslhlnfsvslafitma	779
QY	781	SVYDGKYLPAANMMAGLSLSSSGFGGFTLPCYVILLCRPLDNLSTEHFOASIDQYTRRCGS	840
Db	780	slvgsgylpavavnlaglttllsgfsgyflpkcyvllclcrpelnlteinhgaesldqytrrcgt	839
QY	841	T 841	
Db	840	t 840	
RESULT 8			
AAy45028			
ID	AAy45028 standard; Protein: 840 AA.		
XX	AAy45028;		
AC			
XX			
DT	31-MAY-2000 (first entry)		
XX			
DE	Rat sensory transduction G-protein coupled receptor-B3 variant #3.		
XX			
KW	Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;		
KW	sensory cell; taste receptor cell; screen; taste modulator;		
KW	pharmaceutical; food industry; taste topographic map; tongue; variant		
XX			
OS	Rattus sp.		
OS	Synthetic.		
XX			
FH	Key		
FT	Misc-difference 90		
FT	Location/Qualifiers		
XX	/note="Wild type Ala substituted by Gly"		
XX	W0200006592-AL.		
XX	10-FEB-2000.		
XX	27-JUL-1999; 99WO-US17099.		
XX			

XX 28-JUL-1998; 98US-0094465.
 PR (REGC) UNIV CALIFORNIA.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;
 XX WPI; 2000-205451/18.
 DR
 XX New isolated sensory transduction G-protein coupled receptor. useful
 PT for developing products for use in studying and modulating the taste
 PR transduction pathway -
 XX
 PS Disclosure; Page -: 83pp; English.
 XX

CC The present sequence is a polymorphic variant #3 of rat taste cell
 CC specific G-protein coupled receptor, GPCR-B3 which is involved in
 CC sensory transduction. GPCR-B3 is specifically expressed in foliate
 CC and fungiform cells, with lower expression in circumvallate taste
 CC receptor cells of the tongue. The protein has an extracellular domain,
 CC seven transmembrane domains and an intracellular domain.
 CC The GPCR-B3 sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 CC Note: The present sequence is not given in the specification but is
 CC derived from rat GPCR-B3 sequence shown in page 75 (AAV45021).
 CC
 XX
 SQ Sequence 840 AA;

Query Match 74.7%; Score 3357.5; DB 21; Length 840;
 Best Local Similarity 74.0%; Pred. No. 0;
 Matches 622; Conservative 81; Mismatches 137; Indels 1; Gaps 1;

QY 1 MLCTRRLVGLQLLISCCAFACHSTRESSPDFTLPBDYILAGLFRPHSGCLQYRHRPEVT 60
 DB 1 mlfwaahl-lslqlvycwafscqrlespgfslpgdflglfshgclqlvrhprlvt 59
 QY 61 LCDRSCSFNEHGYHLFQARLGEIEINNSTALLPNTLGLQYLYDCSDANVATLRLVS 120
 DB 60 scdrpsfngfghyflqanrfveelnnsgllpntllgelydvcsaanvayatlrvla 119
 QY 121 LFGQHIEIQLGDLHSPYLAIVIGPDSTNRATTAALLSPFLVPMISYAASETLSVKR 180
 DB 120 lqgrhlelqkdlrhnskvaflgpdntdhavtaallgpfimpvysaasvlsakr 179
 QY 181 QYPSFLRTPNKDYQEVETVLLQKFGWVISVSSDDYQGLQVQALENQAGOCICIA 240
 DB 180 ktpstlrvpsdrhqvewvqllqsfqwwslsgyggqglqyqvaibelaivpica 239
 QY 241 FQDMPFSAOVDERMOCMLRHQAQAGATVAVVVFSSROLARFEFESVLTNLTKGWVAS 300
 DB 240 fkdipfsarvgdprnmqmhlaqartlvvfvfnrlharffrsvvanltgkwwas 299
 QY 301 EYMAISRHTGVPQIORIGMVLGVALQKRAVPGLKAFEBAYARADKEADRPCHKSGWCS 360
 DB 300 edwaistylstvgtlqglvtlqvavqrfvpgljkefesyavraaapacpegswcst 359
 QY 361 NOLCEGCAFMAMHTPKLKAFGSSAYNAYRAVYAAHGLHQLGASGACSGRYPPMO 420
 DB 360 nqlceechtfttrmpcltgaifmsaayryveavahhqlhqlgtseicstgpyppwq 419
 QY 421 LLEQIHKVFLHLKDTVAANDRDLSSNIATAMDNCKMTFTVYGGSTWSPVOLINB 480
 DB 420 llqglvkvflhntevaiddngdtlgyddllawdngpewfletlgsaslsphvldink 479

QY 481 TKIQHGKDNQVPKSVSSDCLEGHORVYTGHHCCPEVCPCGAGFLNKSIDLYRCOPCG 540
 DB 480 tkiqhgknqyvpvsvctldclaghnrvvvgshhccfecvpcceagtlflmsehlhlcqpcg 539
 QY 541 KEEMAPGSGOTCPRTVAVFLALREHNSWVLANTLILLIGTAGLPAWHIDTPVRS 600
 DB 540 teewapkestcfrtrevellawhepislvlanclllllyvtaglfawhfpvrs 599
 QY 601 GGRICFLMLGSLAAGSGSYGFEPTPACLLRQALFALGFTIFSCLTVRSFOLIIF 660
 DB 600 ggricflmlgslvagscsfysfifgeptpacllrqlglfagfifscletlrsfqlvlf 659
 QY 661 KESTKVPTRYHAMVQNGAGLEVMISSAOLICLTWLVTWVPLPAREYQRPPLVMEC 720
 DB 660 kfstkvptrytwagngaglfvsvstwhlllcltwlvmvprpreyqrfphvllec 719
 QY 721 TENNSGFIAPLYNGLLISAFACSYLDKDPENYNEAKVYFSLFNPVSIATFETA 780
 DB 720 tevnsgfiatfahlllslstvcslgkelpenyneakcvfslhntvsvlatfeta 779
 QY 781 SYVDGYLPAANMAGLSISGFGGYFLPKCYVILCRPDLNSTEHFOASIDYPRCS 840
 DB 780 slvgsgylpavnvlaglittlsgfsyflpkcyvillcrpelnstehfsgsidytrcgt 839
 QY 841 T 841
 DB 840 t 840

RESULT 9

AAV45022
 ID AAV45022 standard; Protein; 842 AA.

AC AAV45022;

DT 31-MAY-2000 (first entry)

DE Mouse sensory transduction G-protein coupled receptor-B3.

XX Mouse; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;

KW sensory cell; taste receptor cell; screen; taste modulator;

XX pharmaceutical; food industry; taste topographic map; tongue.

OS Mus sp.

PN WO200006592-A1.

XX 10-FEB-2000.

XX 27-JUL-1999; 99WO-US17099.

XX 28-JUL-1998; 98US-0094465.

PA (REGC) UNIV CALIFORNIA.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;

XX WPI; 2000-205451/18.

DR N-PSDB; AA50744.

XX New isolated sensory transduction G-protein coupled receptor. useful

PT for developing products for use in studying and modulating the taste

CC transduction pathway -

XX Claim 22; Page 75; 83pp; English.

CC The present sequence is a taste cell specific G-protein

CC coupled receptor, GPCR-B3 which is involved in sensory transduction.

CC This sequence was obtained from mouse circumvallate and foliate papillae.

CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with

CC lower expression in circumvallate taste receptor cells of the tongue.

CC The present sequence is used to screen compounds that modulate sensory

CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.

XX Sequence 842 AA;

Query Match 74.5%; Score 3348.5; DB 21; Length 842;
 Best Local Similarity 73.5%; Pred. No. 0;
 Matches 619; Conservative 87; Mismatches 135; Indels 1; Gaps 1;

QY 1 MLCSTAR-LVGLQLLISCCMAPACHSTRESSPDTLPGLYLLAGFLPHSGCLQVHRREV 59
 DB 1 mlfaaahlllslqlavaycwfscqfrespsfipgdfllagflshadclqvhrplv 60
 QY 60 TLCDRSCSFNEGHVLPQAMFLGVEINNSTALLPNTLQVQLVYVCSDSANVATLRLV 119
 DB 61 tscdrsdsfngnyhllqamflveelnnstallpntlqyelyvcsdsnvatclrvp 120
 QY 120 SLPGQHILTEGDLHNSPTVLAVIGPDTNRATTAALLSPFLVPMISYAASTELSVK 179
 DB 121 agqgtghlemgrdlrnhsckvvalipgndthavtaallspflmlyssvssvllsgk 180
 QY 180 ROYSEFLTIRNDKYQVTWVLLQKFGWTWISLVGSSDDYQGLGVQLLENQATQOGICI 239
 DB 181 rkfpstfirtlpsdkyqvavivrlqsfgvvslsvgsydyqglqvgaaleatprglcv 240
 QY 240 AFKIDMPFSQAVGDPRMOCILRHLAQAGATVYVVFSSROLARVPESVTLTNLTKVWVA 299
 DB 241 afkdvpsisagagdpdmrmmrlratrtvvvtfsmhlgavftrsvllanlltgvkwa 300
 QY 300 SEANALSRHITGVPCIORIGVAVLQKRAVPGIKAFEEAYARADKEAPRPHKGSWCS 359
 DB 301 sedwastyltnvpglqglvlgvavlgqvpglkefeesyvqavmgaprtcpeswscg 360
 QY 360 SNOUCRECAQAMHTMPKLFKAFSSSATNATRAYAVAHGHLQGLGACSGACSRERYPW 419
 DB 361 tnglcrecahtwmpelgafmsaayvayavahglhqlgycsgtarqrvypw 420
 QY 420 QLEQIHVHFLHNDYVAFNDNRDPLSSYNIADWNGPKMTFVLSSTWSPQULNIN 479
 DB 421 qlldqiyvvnflhkktrvadkdgdpdygdliawdmngpewtlefvgasaslsphldin 480
 QY 480 ETKIQWAGKDNQVPSKSSDCLBEGHQRVVTGFHHCCECVPCGAGFLNKSNDLYRCQPC 539
 DB 481 ktklqwhgknqgvsvctrdclqghhrvmgshhccfcmpeceagftlnsethltcpc 540
 QY 540 GKEEAPRPGSOTCPRTYVFLALRHTSWLLAANTLLLLLLGTAGLFAHMLDPPVRS 599
 DB 541 gteewapessactsrteflgwhpelsvllaentlllllgtaglfawrlhbpvrs 600
 QY 600 AGGRCLFMLSLAAGSGSLVGFCEPTRPACILROALFALGFTFELSLVRSQOLITI 659
 DB 601 aggrclfmisglvagseslsvfgrkvpacillqpflsfifaifscilstrsqvll 660
 QY 660 FKFSKVPTEYHAWQNHAGLFWMISSAOLILCLTWLVVTPPLPAREYORPHLVMLE 719
 DB 661 fkfstkvpftfhtwagngagifvsvstvhllfcltwlamtprptreyqfplvll 720
 QY 720 CTENISGLFIATFLYNGLISAFACSTLGDLPENYNEAKCVTSLLFNYSWIAFTT 779
 DB 721 ctenvsvglvafahnlisistfvcslgkelpenyeakcvtslllhfvswiaftm 780
 QY 780 ASVYDGLVLPANMAGLSLSSGFGYFLPKCYITLCRPDINSREHQAISODYTRRG 839
 DB 781 sslyggsylpavnvlaagiatlsgfsgyflpkyvllcrlpeinnthqasldqylrrcg 840
 QY 840 ST 841

DB 841 tt 842

RESULT 10
 AAEL1971
 ID AAEL1971 standard; Protein; 366 AA.

XX AAEL1971;

DE 18-DEC-2001 (first entry)

XX Human novel G-protein coupled receptor (NGPCR) protein #3.

KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antisense-therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeat rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease.

OS Homo sapiens.

EH Key Location/Qualifiers
 FT Misc-difference 174
 FT /note= "Encoded by RCA"

XX MO200172842-A2.

XX 04-OCT-2001.

XX 28-MAR-2001; 2001WO-US09996.

XX 28-MAR-2000; 2000US-192978P.

XX (LEI-) LEXICON GENETICS INC.

XX Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
 PI Zambrowicz B, Sands AT;

XX WPI: 2001-616474/71.

XX N-PSDB: AAD19503.

XX Novel isolated polynucleotides encoding human G protein coupled
 PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
 XX
 PS Claim 5; Page 77-78; 80pp; English.

CC The present sequence is human novel G-protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span
 CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeat rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.

XX Sequence 366 AA;

Query Match 40.1%; Score 1803; DB 22; Length 366;
 Best Local Similarity 99.4%; Pred. No. 4e-174;
 Matches 334; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX 08-MAY-2000 (first entry)
 XX Rat GPCR-B4 polypeptide.
 DE Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; rat;
 KW taste transduction pathway; taste receptor; foliate; fungiform; food;
 KM circumvallate; taste signaling; pharmaceutical.
 XX Rattus sp.
 XX WO200006593-A1.
 PN 10-FEB-2000.
 PD 27-JUL-1999; 99WO-US17104.
 PE 28-JUL-1998; 98US-0095464.
 PR 17-DEC-1998; 98US-0112747.
 XX (REGC) UNIV CALIFORNIA.
 PA Zuker CS, Adler JE, Lindemeier J;
 PI WPI; 2000-195257/17.
 DR N-PSDB; AA258963.
 XX New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway and for generating taste topographic maps -
 XX Claim 22; Page 69; 76pp; English.
 XX The invention provides nucleic acids encoding rat, mouse and human
 CC sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
 CC GPCR polypeptides are components of the taste transduction pathway. The
 CC nucleic acids can be used to identify taste cells and as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
 CC for identifying subpopulations of taste receptor cells such as foliate,
 CC fungiform, and circumvallate taste receptor cells. The polypeptides can
 CC be used for identifying compounds that modulate sensory signaling in
 CC sensory cells. Such modulators of taste transduction are useful for
 CC pharmacological and genetic modulation of taste signaling pathways. These
 CC modulatory compounds can then be used in the food and pharmaceutical
 CC industries to customize taste. The present sequence represents a
 CC rat GPCR-B4 polypeptide.
 XX
 SQ Sequence 843 AA;

Query Match 30.5%; Score 1370; DB 21; Length 843;
 Best Local Similarity 37.8%; Pred. No. 1.6e-129;
 Matches 318; Conservative 145; Mismatches 316; Indels 62; Gaps 22;

QY 31 DFTLPGVYLAGLFFPLSLGCGVQVHRP--EYTLDRSGSFHEHGYHLFQARLGEVEELNN 88
 DB 29 dfhlagvyllyglltllhanvksishlylqypkcn-eflmkvlygnlmqamrtfaveeln 87
 QY 89 STALLPNTLTGYQLDYVCSDSANVYATLRVLSLPGOHHEIQGLD-----HYSPVL 141
 DB 88 csalllpgvlllygemvdcyismnh-----pglyflaeddallpalkdysqymphv 139
 QY 142 AVIGPDSTNRATTAALISPELVPMISYAAASEFLSVKQYPSFLRTIPNDKYOVETWVL 201
 DB 140 avigpdseesaitvsnllshlllpqlytsaisdklrdrkhrfpmrlrtvpsathleamvq 199
 QY 202 LLOKFGWTWISLVGSSDDYQGLQVQALENAT--GGGTICIAKDIMP--SAQV--GDERM 256
 DB 200 lmvhfgmwvlyvvsdddygvenshllsqrltktsdldiclaqdevlipessqvmrseqr 259
 QY 257 QC--LMRHIAOGATVVVVSRRQLARVFESVVLTLNLGKVVWVASEAMALSRHITGVPG 314

DB 260 qldmlklrlrtsarvvvvpelslysfhnevlwntfgtvlaseswaidpvlhnlte 319
 QY 315 IQRIGWVIGVAIQKRAVGLKAFFENAVARADKEA-PRCHKGSGWSSQQLRECOAFMAH 373
 DB 320 lrlhgtfllygvltlqrvslpfgsqfr---vrickpypvptnltlcttcqddac----- 370
 QY 374 TMPRLKAFS-----MSS---AYNARYAVVAHGHQHLGSCASGASRCGRVPMOLLEQIH 426
 DB 371 -lntksfnnlllsgervvysvsaayavahalnrlggcnrvctckqkypwqllreilw 429
 QY 427 KVHFLHKDVAFNDNRDPLSSYNNIAMDWNGPKWTFVLGSSWSPVOLNINE-TKIQW 485
 DB 430 hvnftllgnrlffddqgdmplldliqgwldlsqnpfgsias--yspskrltylnvsw 487
 QY 486 HGKDNQVPSKVSQSSDCEGHQRYVTGFHNCFECPGACAGFLNKS-DLYRCQPGKEEW 544
 DB 488 ylpnltvpsmcsksqpgqmkksvglhpcctfecldcmpgtylnrsadefnclscpgsmw 547
 QY 545 APEGSOCTCFPRVVFALREHSMV--LLAA-----NTLLLLLTGTAFLAMHLDTPYVR 598
 DB 548 sykndltcftgrrptflwhevpplvallaalgfstlaillf-----lfrhfgtpmvr 601
 QY 599 SAGGRCLFLMLGSLAAGSGSLYGEFGEPTRPACLRLQALFALGFTPLSCLTVRSFQDII 658
 DB 602 saggmctflmlvplllatgmvpyvgpptvscfcrqgaftefcfscitcstlvsfqlvc 661
 QY 659 IFKSTKVPFTYHAMVQNHGAGLFPYMISAAQOLICLTWLVVWPLP-AREYRPHLYM 717
 DB 662 vfknaarlpysagslfmryhpgyvaftaikaivaivgnmlatlnlpgirdpddpnlli 721
 QY 718 LECTETNSLGFILAFVLYVGLSISAFACSYLQKDPENYNEAKCTPSILTFNFSWIAEF 777
 DB 722 lscbpnyrnqllfntsmlllsvlgfstaymgkeiplynakeliltstfssislcl 781
 QY 778 TTASVYDQKYVPAANMMAG-LSSLSGFGGYFLPKCYVILCRPDLNSTEHFQASIODYTR 836
 DB 782 tfmsvhdgvltlmdllvltvlnflaigl-gylgpkcymlltypentasayfnsmlqylm 840
 QY 837 R 837
 DB 841 r 841

RESULT 13
 AA77557
 ID AA77557 standard; Protein; 843 AA.
 XX
 AC AA77557;
 XX
 DT 08-MAY-2000 (first entry)
 XX
 DE Mouse GPCR-B4 polypeptide.
 XX
 KW Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; mouse;
 KW taste transduction pathway; taste receptor; foliate; fungiform; food;
 KW circumvallate; taste signaling; pharmaceutical.
 XX Mus sp.
 XX WO200006593-A1.
 PN 10-FEB-2000.
 PD 27-JUL-1999; 99WO-US17104.
 PE 28-JUL-1998; 98US-0095464.
 PR 17-DEC-1998; 98US-0112747.
 XX (REGC) UNIV CALIFORNIA.
 PA Zuker CS, Adler JE, Lindemeier J;
 PI

DR WPI: 2000-195257/17.
 DR N-PSDB: AA258964.
 PT New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway and for generating taste topographic maps -
 XX
 XX
 PS Claim 22: Page 69-70; 76pp; English.
 CC The invention provides nucleic acids encoding rat, mouse and human
 CC sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
 CC GPCR polypeptides are components of the taste transduction pathway. The
 CC nucleic acids can be used to identify taste cells and as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centers in the brain. GPCR-84 is useful as a nucleic acid probe
 CC for identifying subpopulations of taste receptor cells such as foliate,
 CC fungiform, and circumvallate taste receptor cells. The polypeptides can
 CC be used for identifying compounds that modulate sensory signaling in
 CC sensory cells. Such modulators of taste transduction are useful for
 CC pharmacological and genetic modulation of taste signaling pathways. These
 CC modulatory compounds can then be used in the food and pharmaceutical
 CC industries to customize taste. The present sequence represents a
 CC mouse GPCR-84 polypeptide.
 XX
 XX Sequence 843 AA:
 SQ
 Query Match 30.0%; Score 1350; DB 21; Length 843;
 Best Local Similarity 37.2%; Pred. No. 1.7e-127;
 Matches 311; Conservative 144; Mismatches 329; Indels 52; Gaps 18;
 QY 31 DFTLPDGYLAGLPHSLGCLQYRHRP--EVTLCDRSCSENEH-----GYHLEQANRLGV 83
 DB 29 dftlpadgyllgglftlhanvksvshlsylyqpkc-----neynmkvlygnlqnamrfav 82
 QY 84 EEIINSTALLPNTLTGOLYDVCSDSANVYATRLVLS-----LPGQHIELOGLDLHYSR 138
 DB 83 eeiinssallpnyllygmwvdcysnllqpylyflsqldfip-----llkdyqytp 136
 QY 139 TYLAVIGPSTNRATTAALISPELPMISYAASSSETLSVKROYPSFLRTIPNDKQVEN 198
 DB 137 tyvavigpnsesatvsnllsyflvpyvtysatldkldkrrfpmlltvsathiea 196
 QY 199 MYLLLOKFGWTVSLVSSDDYGOGLVQALENGATGOG--ICIAFKDIMPE---SAOVD 253
 DB 197 myqlmwthfgwnvlyvlsdddygreshllsqllntgclafgevlpypepqnqvrpe 256
 QY 254 ERHQC--LMRHLAGATVAVVSSRQLARVFESVVLNLTCKVAVASAMLSRHNG 311
 DB 257 eqdqalldkrlrrtsarvavvlspelslhnffrevlvmfvgfvaasesvaldvlm 316
 QY 312 VPGIORIGWLVGLVQIRAVPGKAEFEAVARADKEAPRCHGKSSCNSQLRRCQAFM 371
 DB 317 lcelrhtgftlgtvgrvtsfgrvhdhkgymmettsrltc--ngdcdaemlt 374
 QY 372 AHTMPKIAFSMSSAYNAVAVYAVAHGLHQLGCASGRGVYPMOLLQIHVHEL 431
 DB 375 esfnvmlsgervvysavayavahlhrlhcnqyrcdkgyvpyqlreihwnft 434
 QY 432 LHKDYAFANDNDPLSSYNIAMDMNGPKWTFVLCSSWSPQOLINE--TKIQHGKDN 490
 DB 435 llnqglftdegdmpmllldlqwgqlsqnplfqsias--yspeltreltlylsnswytlm 492
 QY 491 OYPRKSSDCLLEGHGVVTVGFHHCFCFCVPCGAGTFLNKS--DLVYCOGCKREMPG 549
 DB 493 tlpismcsccpgqmkkrpighpcfecvdeppdylnsvdelrnclosgmswmykm 552
 QY 550 QYCFPRVYFLALREHTSWV--LLAA-----NTLLLLLLGTAQLFAMHLDTPVRSAGR 603
 DB 553 laccfrlafllewhvevplvtcllaagfistlaill-----lfvrhgtqtmvrsagrp 606
 QY 604 LCFIMGLSLAAGSGSLYGFGEPTRPACLLRQALFALGFTIFLSCLTVNSFOLLITFKPS 663

DB 607 mcfmlmlyp1llatfmgwpyvypplvtfscfgrqgftfvcfscvclstlvsfsgivcfkma 666
 QY 664 TKVPFTYHAMVQNHGAGLVMISSAOLLICLWLWVWPLP-ARYQRPPIHVMCEH 722
 DB 667 rrlpsaygftwmtyhgpvvtavfllavkaavlvaqumlatlmpdgrddpdpnlliscnp 726
 QY 723 TNSIGFILAFLVNGLISAFACSYLGDLPENYNEAKCVTESLLEPNYSWIAFFTASV 782
 DB 727 nyrngllfntsmldllsvlgfaygkelpnynaeakflltsmtsfsslstcfmsv 786
 QY 783 YDGKRYPANMAG-LSSISGCGGFLPKCYIILCRPDLNSENHQASIDQYTR 837
 DB 787 hdgvlvtlmdllvtvinflaigl-gyfpbkoymllyperntsayfnsmlqgytmr 841
 RESULT 14
 AAEL1972
 ID AAEL1972 standard; Protein: 234 AA.
 XX
 AC AAEL1972;
 XX
 DT 18-DEC-2001 (first entry)
 DE Human novel G-protein coupled receptor (NGPCR) protein #4.
 XX
 KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antisense-therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeate rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease.
 XX
 OS Homo sapiens.
 PN WO200172842-A2.
 XX
 PD 04-OCT-2001.
 XX
 PE 28-MAR-2001: 2001WO-US09996.
 XX
 PR 28-MAR-2000: 2000US-192978P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
 PI Zambrowsicz B, Sands AT;
 XX
 DR WPI: 2001-616474/71.
 N-PSDB: AAD19504.
 XX
 PS Novel isolated polynucleotides encoding human G protein coupled
 PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
 XX
 PS Disclosure: Page 78-79; 80pp; English.
 XX
 CC The present sequence is human novel G-protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span
 CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeate rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.

XX Sequence 234 AA;
SO
Query Match 27.6%; Score 1238; DB 22; Length 234;
Best Local Similarity 100.0%; Pred. No. 5,2e-117;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 608 MLGSLAAGSSGLYGFGEPTPRACILRLRALFALGTTIFLSCITVRSFOLIIIFKSTKVP 667
D 1 mlgsiaagsgsllygfgeptrpacilrallgltflscitlvrsfqlilifkfstkvp 60
QY 668 TFYAWVONHAGLFVWISSAQAOLICTWLVWTPLPAREXORPHVMECTETNSLG 727
D 61 tfyawnvnhagglfwmssaaqlilcltlwvtplpareyqrphlvmlectetnslg 120
QY 728 FILAFLYNGLLISAFACSYLGKDLPENYNEAKCVTFSLFNFVSWIAFFTASVYDKY 787
D 121 filaflyngllisafacsylgkdlpenyneakcvtfslfnfvswiafftasydydky 180
QY 788 LPAANMAGLSLSGFGGYFLPKCYVILCRPDLNSTHFQASIODYTRRGST 841
D 181 lpaanmaglsissgfygylpkcyvllcrpdlnstehfgasldytrrcgst 234
RESULT 15
AAE10371
ID AAE10371 standard; Protein: 858 AA.
XX
AC AAE10371;
XX
DT 10-DEC-2001 (first entry)
XX
DE Rat taste receptor, rtrr3 protein.
XX
KW Rat; taste-cell-specific G protein-coupled receptor; rtrr3; drug;
genetic modulation; pharmaceutical; taste sensation; food industry;
chemosensory transduction.
XX
OS Rattus sp.
XX
PN W0200166563-A2.
XX
PD 13-SEP-2001.
XX
PE 07-MAR-2001; 2001WO-US07265.
XX
PR 07-MAR-2000; 2000US-0187546.
PR 07-APR-2000; 2000US-0195536.
PR 06-JUN-2000; 2000US-0209840.
PR 23-JUN-2000; 2000US-0214213.
PR 17-AUG-2000; 2000US-0226448.
PR 03-JAN-2001; 2001US-0239227.
XX
PA (SENO-) SENOMYX INC.
XX
PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
XX
DR WPI: 2001-582267/65.
DR N-PSDB: AAD17514.
XX
PT New mammalian taste-cell-specific G protein-coupled receptor
polyptides for identifying compounds that modulate taste signaling
are useful in food, to modulate the sweet taste of foods or drugs -
XX
PS Claim 152; Page 77; 119pp; English.
XX
CC The invention relates to mammalian taste-cell-specific G protein-coupled
receptors, rtrr and their corresponding cDNA molecules. Taste receptors,
rtrr are useful for screening compounds which are used to activate or
modulate chemosensory transduction, such as taste sensation. The
CC identification and isolation of novel taste receptors and taste
signalling molecules allow for new methods of chemical and genetic

CC modulation of taste transduction pathways. The taste modulating
CC compounds are useful in pharmaceuticals and food industries to improve
CC the taste of a variety of consumer products, or to block undesirable
CC tastes, e.g., in certain pharmaceuticals. rtrr's are also useful in
CC biochemical assay for identifying tastant (rtrr) ligands having binding
CC specificity for rtrr involved in taste signalling. The present sequence is
CC rat taste-cell-specific G protein-coupled receptor, rtrr3 protein.
XX
SO Sequence 858 AA;
Query Match 26.0%; Score 1170; DB 22; Length 858;
Best Local Similarity 33.3%; Pred. No. 3.5e-109;
Matches 277; Conservative 142; Mismatches 358; Indels 54; Gaps 17;
QY 29 SPDFTLPGDYLAGLFPPL---HSGCQVRHREPVTLCDRSGCFNHHGTHLFOAMLYGEE 85
D 26 sqdfkagdyllgglfpjteeatlmgtrqpnllctr---fspjgflflamamkmavee 82
QY 86 INNSTALLPNTLGYQVLDVDCSDS-ANYATLRVLSLPGQHHELOGDLHSPVLAIV 144
D 83 inngsallpnlrlygdlldtcepevntmkpslmakvsgsaacyntgqprvlayl 142
QY 145 GPDSTNRATTPAALLSPFLVPMISYASSETLSVRKQYFRTIPNDKYQETWVLLQ 204
D 143 gphselaligkffsfllmpqvsaaendrdsdetrfsfrrtvsdrvdqavvllq 202
QY 205 KFGWTWISLVSSDDYDGLGVQALENQATGOCICIAFKDIMEFSQVQDE--RMOCIMRH 262
D 203 nfwvmaalgsdddygrelsfagslrsrgiciahneglvphqhtsqqglgkvvdvrlq 262
QY 263 LAQAGATVWVWSSRQRLARVFESVLTNLCKVWVAVSAMLSHRITGPIORIGWL 322
D 263 vngskvqvavvlfaasavyslfsyailndlskvwvaesvltldmlrldarvavl 322
QY 323 GVAIDKRAVPGIKAFEEAVARADKEAPPCCHKSGWSS-----NOLCRECOAF 370
D 323 gflrgallpelsyvetrlal---aadr-----ftcaalkaeldleervmprscqdy 375
QY 371 MAHT-----MPKLAFSM-SSAYNAYRAVYAVAHGHLQGCASACSGR- VYRWOLLE 423
D 376 mlgnlssgldmqnlssagqlnhqifaayaavysaqaalhtlqcnvachltsevpqbwjle 435
QY 424 OIHKVFLLHKDTVAFNNRDLPLSSYNTIAMQNPMTFVILGSSVSPVOLNINERKI 483
D 436 mmymslrrardltlqfdakgsvdmeaydlkmwvgsbtpvllhtvglnfgl---lqghskm 492
QY 484 OMHGKDNQVPKSCSSDCLGHRVVTGFHHCFCFCVPCGAGTFLNKSDLYRCOPGKRE 543
D 493 ywpg--ngvpyvsgsrqckdgqrvrvkghfscycvcdckagyskhhpddftcpcqkdx 550
QY 544 WAPESQTCFPRVTVFLAREHTSWVLAANTLLLLILGTRGLRWMHLDPVVSAGR 603
D 551 wpekstctclprpkflawgepavlllllclvlgltlaalgltlvhwdsplvasgys 610
QY 604 L-CF--LMGSLAAGSSGLYGFGEPTPRACILRLRALFALGTTIFLSCITVRSFOLIIIF 660
D 611 lfcrllclglfcl---svllfpgrprsaasciaqgmahplrtgclstclfdgaaelfv-- 665
QY 661 KFTKVPFTFYAWVONHAG---LFVWISSAQAOLICTWLVWTPLPAREXORPHV 716
D 666 --eselpswanwlsyrlgprawlvvlatlveaalcaewylmafpvevtdwylpvev 723
QY 717 MECTETNSLGLTAFVLYNGLLISAFACSYLGKDLPENYNEAKCVTFSLFNFVSWIAF 776
D 724 lehcimrswslglvhltnavlatlclfgltfivgsqprgrnargrlftamlaylwsf 783
QY 777 FTASVYDQKYLPAANMAGLSLSGFGGYFLPKCYVILCRPDLNSTHF 827
D 784 vpllanvgyavpavmgallfcalglilatflhpkcyvllwlpelntgeff 834

Tue Jun 11 13:05:58 2002

us-09-819-946-2.rag

Page 14

Search completed: June 11, 2002, 10:17:41
Job time: 3813 sec

```

|||||
Db 262 CTTCTCCGCGGACCTCAACAGCAGACGACCTTCAGGCTTCATTCAGGACTACAGC 203
|||||
Qy 2506 aggcgtcgcgcctccacctga 2526
|||||
Db 202 AGGCGCTGCGGCTCCACCTGA 182
|||||

RESULT 3
AA853967/c 496 bp mRNA linear EST 31-DEC-1998
LOCUS 311010.1 Soares_testis.NHT Homo sapiens CDNA clone IMAGE:1393866
DEFINITION 3' similar to SM:CSR.BAT.P48442 EXTRACELLULAR CALCIUM-SENSING
RECEPTOR PRECURSOR ;, mRNA sequence.
ACCESSION AA853967
VERSION AA853967.1 GI:2941505
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 496)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the J.M.A.G.E. Consortium/LNLW at:
www-bio.linnl.gov/dbfp/image/image.html

Possible reversed clone: similarity on wrong strand
Insert length: 878 Std Error: 0.00
Seq primer: -40ml3 fwd. ER from Amersham
High quality sequence stop: 223.
Location/Qualifiers
1. 496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1393866"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand CDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTCACCAATCTGAAGTGGGCGCGCCCAATTTTCTTTTCTTTT 3'].
Double-stranded CDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 102 a 143 c 148 g 100 t 1 others
ORIGIN

Query Match 11.0%; Score 278.6; DB 9; Length 496;
Best Local Similarity 95.6%; Pred. No. 2.8e-55;
Matches 306; Conservative 1; Mismatches 11; Indels 2; Gaps 2;
Qy 2207 ggcctctccacagctcgtcgtcgcagctacgtgtaagagcttcagagaaact 2266
|||||
Db 495 GGCCTCTCCATCATGATGCTTCCCTGACGCTACCTGGTAAAGACTTGGACAGAACT 436
|||||
Qy 2267 acaacagagcaaatggtacacctcagcgtcgtctcaactcgtctcgtgacgcct 2326
|||||
Db 435 ACAACAGAGGCCAAATGTCTCACTTCAGCTGCTCAATTCGTCTGATCGGCT 376
|||||

```

```

Qy 2327 tcttaccacagccagcgtctcagcagcgaagtactcgtcgcgcacaaatgatgctg 2386
|||||
Db 375 TCTTCACCAAGCGGCAGGCTCTACAGCGCAAGTA-ANCTTGGGCGCAACAGATGAGCT- 316
|||||
Qy 2387 ggcctagcagcctgagcagcgcgtcgtggtatcttctgctaaagtcagtcaccc 2446
|||||
Db 317 GGCCTAGCAGCCTTGAGCAGCGGCTTCGGTGGGTATTTCTGCTTAAGTCTACGATCC 258
|||||
Qy 2447 tctgcgcgccagacacctcaacagacagagcacttcagcgtcctcaatgaagatacaga 2506
|||||
Db 257 TCTGCCGCCACAGACCTCAACAGCAGACGACTTCAGGCTTCATTCAGGACTACAGA 198
|||||
Qy 2507 ggcgtcgcgcctccacctga 2526
|||||
Db 197 GGCCTGCGGCTCCACCTGA 178
|||||

RESULT 4
BB618551/c 666 bp mRNA linear EST 26-OCT-2001
LOCUS BB618551 RIKEN full-length enriched, 8 days embryo Mus musculus
DEFINITION CDNA clone 5730408M21.5', mRNA sequence.
ACCESSION BB618551
VERSION BB618551.1 GI:16458094
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 666)
AUTHORS Arakawa,T., Carinini,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M., and Hayashizaki,Y.
JOURNAL RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
COMMENT Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carinini,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Iwata,M., Ohara,E.,
Watanabe,M., Yonekura,Y., Ishikawa,T., Ozawa,Y., Tanaka,T., Matsura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carinini,P., Sugahara
,Y., and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamakawa,I., Atzawa
,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

```



```

Db 249 ----- 248
Oy 2154 tgagtgcacagagaccactccctgggtctactgtgctctctctacatgagctct 2213
Db 249 ----- ACAACATCTCTCT 261

Oy 2214 ctccatcagtccttgcttgctcagctacgtgtaaggacttgcacagaaactacaaga 2273
Db 262 CTCATCATGACACCTTTGCTCTGACGTAACCTGGTAAGAACTGCCGGAACATAAAGA 321

Oy 2274 ggcacaatggtcactcagctcagctctcctaactcgtctcctgagatgcctctac 2333
Db 322 AGCCAAATGTCTACCTTCACCTGCTGCTCTCTCTATCTATCTGATCGCTTCTCTAC 381

Oy 2334 cagggcagaggtctacagagcaagtaactgctgcggcacaatagatgctggtgag 2393
Db 382 CATGTCACGACATTTACCGGGCAGCTACTACCGCGGTCAATGCTGCGAGGGCTGAC 441

Oy 2394 cagcctgagcagcgctcgtggtggtatctctgctcagtagtctgctgctgcg 2453
Db 442 CACTCTGATGCGCGGCTTCACGCGGCTATTCTCTCCCTAAATGCTATGATTTCTGCGG 501

Oy 2454 cccagacctcaacagcagagca 2477
Db 502 TCAGAACTCAACACACAGACA 525

RESULT 6
CNS033X/c 1014 bp DNA linear GSS 15-MAY-2000
DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone
209N08 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL226735
VERSION AL226735.1 GI:7885667
KEYWORDS GSS: genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetradon.
REFERENCE 1 (bases 1 to 1014)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1014)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1014)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
REFERENCE This sequence is a single read and was generated as part of a large
COMMENT genome clone-end sequencing project of the Tetradon nigroviridis
http://www.genoscope.cns.fr/Tetradon.
FEATURES
Source 1..1014
/organism="Tetradon nigroviridis"
/db_xref="taxon:99883"
/clone="209N08"
/clone_lib="G"
/note="Genoscope sequence ID : COAG209D04L1-end : T7"
BASE COUNT 283 a 209 c 322 g 196 t 4 others
ORIGIN

```

```

Query Match 8.2% Score 208; DB 12; Length 1014;
Best Local Similarity 53.2%; Pred. No. 2e-38;
Matches 478; Conservative 3; Mismatches 414; Indels 4; Gaps 2;

Oy 1594 gaccctcaagatgacagcctgtgtggaagaaggtggtgacctgaggaagcagacc 1653
Db 896 GATCCCTACAGCTGGCTCCCTCGTATAGAACACACAGCTGTCCACAGCTGGAAAGACCTCC 837

Oy 1654 tgcctccgagactggtggtgttcttgcttgcttgctgagcaaccttg-gtgcgtc 1712
Db 836 TKTCTCCACGAGAGGTGAGTGGGTCCGTTTACACAGACCCCGCGGTGTGTATCTCT 777

Oy 1713 ggcagataacagcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 1772
Db 776 GCGCCGGGCGCTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 717

Oy 1773 cctagacacccctggtggtgaggtcagaggggagcgcgtgcttcttctatgctgc 1832
Db 716 CTACAAACACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 657

Oy 1833 cctgagcaggtgagtgagcagctctatgctctcttggtggaaacacagagcgtcgt 1892
Db 656 CCTCAGCCTCTGTATACATCAGCGGTCTCTTTACTTGTGAGACCGCATGAAAGCTTTTG 597

Oy 1893 ctgtgtaagcagagccctcttgcccttggttcaacatctctgctcctgctgaagt 1952
Db 596 TATCTTGAGGTTTATGCGCTTCTTTTATCTTACGCTGTGTGTGTGTGTGTGTGTGTGT 537

Oy 1953 tgcgtcatcacaatcatcatctcaagttcttccacaaggtacatcttaca 2012
Db 536 CCGCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 477

Oy 2013 cgcctggtgccaaacacagcgtgcgtgcgtgtgtgtgatgatacagcagcagc 2072
Db 476 CTGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 417

Oy 2073 gctatcgtcctaactgctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2132
Db 416 AGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 357

Oy 2133 ctcccccatcgtggtgctgtggtgacagagacacactccctggtctcatactg 2192
Db 356 CTACCCAGACAAATACATCCTGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 300

Oy 2193 ctctctcaaatgctcctctctcatcagtgctgtccttgctgcagctacgtggtgaaga 2252
Db 299 ATTTTGTGCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240

Oy 2253 ctggcagagaactacaacagagcaaatggtcactcctcagcctgctcttaactg 2312
Db 239 CTTCCCGAAGATCAACAGAGGCGCAAGCCATAACTTCTGCTTCTGCTCTATCT 180

Oy 2313 gtcctgatacgcctctctcacaacagcagcagcgtctcagagcaagtaactgctgc 2372
Db 179 CACCTGATCATATTTGGCACCGGCTTCAATGCTTCAACAGGCAATATCATCACACCT 120

Oy 2373 caacatgagctggtgctgagcagcctgagcagcgtctggtggtatcttctcctaa 2432
Db 119 CAACCCCTGCGCGGTCTCTCCAGCGCTTACCTTCTCTGCTGTGCTTCTCTCCCA 60

Oy 2433 ggtcagtgatcctctcgtgcgcagcactcaacagcagcagcagcagcagcagc 2491
Db 59 ATGTTACATTTATCATCTTTCACGCTCAAGAACACCCAGAGATGCTTCCAACTTCCA 1

RESULT 7
A1415100/c 421 bp mRNA linear EST 09-FEB-1999
LOCUS A1415100/c
DEFINITION mb98d01.x1 Soares mouse p31MF19.5 mus musculus cDNA clone
IMAGE:337441 3' similar to SW:CA8R_EAT P48442 EXTRACELLULAR
CALCIUM-SENSING RECEPTOR PRECURSOR ;, mRNA sequence.

```

Accession	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal	Comment
AI415100	AI415100.1	GI:4258604		house mouse.					
				house mouse.					
				Mus musculus					
				Embryoto: Euthera: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.					
				1 (bases 1 to 421)					
				Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Thelsting,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.					
				The Washu-NCI Mouse EST Project 1999					
				Unpublished (1999)					
				Contact: Marra M/Washu-NCI Mouse EST Project 1999					
				Washington University School of Medicine					
				4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA					
				Tel: 314 286 1800					
				Fax: 314 286 1810					
				Email: mouseest@wustl.wustl.edu					
				This clone is available royalty-free through LNL; contact the					
				IMAGE Consortium (info@image.llnl.gov) for further information.					
				This clone was previously sequenced on the 5' end only, this new					
				data is from the 3' end					
				Possible reversed clone: similarity on wrong strand					
				High quality sequence stop: 373.					
				Location/Qualifiers					
				1..421					
				/organism="Mus musculus"					
				/db_xref="taxon:10090"					
				/clone="IMAGE:337441"					
				/clone_lib="Soares mouse p3NMF19.5"					
				/dev_stage="19.5 dpc total fetus"					
				/lab_host="DH10B (ampicillin resistant)"					
				/note="Vector: pUTR3 (pharmacia) with a modified					
				polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA					
				was primed with a Not I - oligo(dT) primer [5']					
				TGTTACCAATCTGAAGTGGAGCGCGCCATTTTTTTTTTTTTTTT 3']					
				double-stranded cDNA was size selected, ligated to Eco RI					
				adapters (Pharmacia), digested with Not I and cloned into					
				the Not I and Eco RI sites of a modified pUTR3 vector					
				(Pharmacia). Library went through one round of					
				normalization to a Cot = 5. Library constructed by Bento					
				Soares and M.Patima Bonaldo. RNA was kindly provided by					
				Dr. Minoru Ko (Wayne State University)."					
				BASE COUNT 90 a 117 c 114 g 100 t					
				ORIGIN					
				Query Match 7.8%; Score 196.4; DB 9; Length 421;					
				Best Local Similarity 83.7%; Pred. No. 7.3e-36;					
				Matches 221; Conservative 1; Mismatches 42; Indels 0; Gaps 0;					
				2263 aactcaacagagcgcaaatggtacgtctcaacgtctctcaacttcgttccttgatc 2322					
				421 AACCTATACGAGCAACAAATGTCTACCTTCACCCCTGCTCTCCACTTCGTATCTGGATC 362					
				2323 gccctctccacacagcgccagcgtctcaagcagcaagtactcgtcgtcgccacaatgaty 2382					
				361 GCTTCTTCACACATGTCCAGCATTTTACCAGSGGACACTACTACCCGGGTCAATGTGCTG 302					
				2383 gctggcgtgagcagcctgagcagcggtctcggttggtattcttcgtctaagtgtactgty 2442					

RESULT	8
AZ994369/c	
LOCUS	
DEFINITION	AZ994369 626 bp DNA linear GSS 27-APR-2001
ACCESSION	2M0279L18R Mouse 10kb plasmid U0GC2M library Mus musculus genomic
VERSION	AZ994369
KEYWORDS	AZ994369.1 GI:13865596
SOURCE	GSS.
ORGANISM	house mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 626) Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tinge,A., von Niederhausern,A. and Wright,D., Weiss,R. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
TITLE	Unpublished (2000)
JOURNAL	Contact: Robert B. Weiss University of Utah Genome Center
COMMENT	Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel.: 801 585 5606 Fax: 801 585 7177 Email: ddunne@genetics.utah.edu Insert length: 10000 Std Error: 0.00 Plate: 0279 Row: L column: 18 Seq primer: CACACAGGAAACACTATGACC Class: plasmid ends High quality sequence stop: 626. Location/Qualifiers 1..626 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="U0GC2M0279L18" /clone_id="Mouse 10kb plasmid U0GC2M library" /sex="Female" /lab_host="E. coli strain XL10-Gold, F1-resistant, F-" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g114732114[g]A[129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT	142 a 183 c 177 g 124 t
ORIGIN	
Query Match	7.7%; Score 194.4; DB 12; Length 626;
Best Local Similarity	83.6%; Pred. No. 2.6e-35;
Matches 219; Conservative 1; Mismatches 42; Indels 0; Gaps 0;	
ctacaagaggccaatyygtcaaccttcaagcctcttcaacttcgtctgcattgcgc 2324	

DB	Accession	Source	Organism	Reference Authors	Title Journal	Comment
Db	626	CTATNACGAAGCAAAATGTGTACACTTACGCTGTCTCTCCACTTCGATTCCTGGATGCG	567			
Qy	2325	ctcttcacacagcgcaagcgtctctacagcagaagtaactcgtctcgccagcaacatgatgc	2384			
Db	566	TTTCTTTCACCAATGTCTACCACTTTTACACAGGAGAGCTACCTACCCGGGTGATGTGCTGGC	507			
Qy	2385	tggagcttgacagcagctcgacagcagcgtctcggtcggtgattcttcgtccaaagtactgat	2444			
Db	506	AGGGCTGGCCACTCTGATGATGGCGGCTTACGGCGCTATTCTCTCCCTTAATGCTACGTAT	447			
Qy	2445	cccttcgccccagagcacttcacacagacagcagcagcacttcaggccttcacaggacatac	2504			
Db	446	TCTCTGCGGTGACAGAACTCAACAACACAGAACTTTTACGCGCTTCATTCACAGACTACAC	387			
Qy	2505	gagggcgctggcggtctccactga	2526			
Db	386	GAGGCGCTGGCGCACTACTCTGA	365			
RESULT	9					
LOCUS	A1562167/c					
DEFINITION	A1562167	298 bp	MRNA	linear	EST 25-MAR-1999	
ACCESSION	A1562167					
VERSION	A1562167.1					
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 298)					
	Marra, M., Hallier, L., Kueba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Streptoe, M., Thaising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rilter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.					
	The WashU-NCI Mouse EST Project 1999					
	Unpublished (1999)					
	Contact: Marra, M/WashU-NCI Mouse EST Project 1999					
	Washington University School of Medicine					
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA					
	Tel: 314 286 1800					
	Fax: 314 286 1810					
	Email: mouseest@wustl.edu					
	This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.					
	MG:663147					
	This clone was previously sequenced on the 5' end only, this new data is from the 3' end					
	Possible reversed clone: similarity on wrong strand					
	Possible reversed clone: POLYT not found					
	High quality sequence stop: 286.					
FEATURES						
SOURCE						
	1..298					
	/organism="Mus musculus"					
	/strain="NIH/Swiss"					
	/db_xref="taxon:10090"					
	/clone="IMAGE:1260595"					
	/clone_lib="Stratagene mouse heart (#937316)"					
	/sex="pooled"					
	/tissue_type="heart"					
	/dev_stage="13 day embryos"					
	/lab_host="SOLR (kanamycin resistant)"					
	/note="Organ: heart; Vector: pBluescript SK-; Site: 1; EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer: Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GATTTGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGCTTTTTTTTTTTTTTTTTT 3'."					
BASE COUNT	59 a	101 c	87 g	51 t		
ORIGIN						

[illegible]

ORIGIN

Query Match 4.9%; Score 124.2; DB 12; Length 973;
Best Local Similarity 51.4%; Pred. No. 1.3e-18;
Matches 361; Conservative 4; Mismatches 322; Indels 16; Gaps 3;

```
QY 1626 agatgggacactgagggaagcagacgtgtcccgccactggyggttttggcctt 1685
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 AGAGGGCTCCAAAGGAACTCCAGCTGGCTCCAAACAACTGCTTCTCTCTG 338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1686 gcgtagacacacccctctggtgctgctgagctaacacgctgctgctgctgct 1745
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GAAGGACACCTTTGGGCGTTTCATGCGCTTCCACCTGGGCACTCCTCTGCT 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1746 tggagctgctgctgctgttgcctgagcaaacccctggtgtagagtcagcagg 1805
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GTWGAACATGGCTGCTTCTGTACAAAGCGGACACTCCGGTGGTGAAGCAGCAG 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1806 cgcgctgctgtcttctttagctggtcctgagcagcaggtagtgagcctctatgctt 1865
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 ACCCTTAGCCCAAGCCATCTCTTCTCTCTGCTATCATCATCATCAGCCGCGCTGT 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1866 ctltggggaaccacaagcctgctgctgctgctgctgctgctgctgctgctt 1925
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 TGTGGGCGAGCMACAGACCTCCAGTGCAGAGCCCGCAGGTCCTCGGCATCAGCTT 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1926 caccatctctgctgctgctgacagtcgctcatctcaactaatcatctcaagtt 1985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 CACTCTGTGTATCTGTCTGTCATGTCAGACTTTCAGATCTGTAGCTTCTCAGTT 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1986 ttccaccagaagttacattctacacgctggttccaaaaccacggtgctgctgtt 2045
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CAACCCCGGGCTGC-----AGGACGTGCTGAGAGATCTATCGGCCGTACGCCAT 689
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2046 tgtatgatacagctcagcgcgcgctgcttatactgctgctgctgctgctgctgac 2105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 690 CATCACCGCTGTGCGTGTGTCAGACGTCATCTGCTGCTGCTGCTGCTGCTGCTG 749
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2106 cccacgtcctgctgaggaataccagcgtctccccaatctggtgtagctgagtgacaga 2165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 750 CCCGTACATACACCATCGGGCAAGCAGCACCTCTTTGATG-----CACTGGCACGA 803
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2166 gaccacactccctggtgcttcatatgctgcttctctacaatgagcctctccatcagtc 2225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 804 AAACCTGACGTGCGCTTCGGGGTATGCTGGGCTACATCGCCCTCTGCGCTTTGTGT 863
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2226 -cttgcctgagctacctggttaagaaactgcccagaaactacaacgagcgaatgyg 2284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 864 GCTTCGTCGTTCCTTCAAGGSSCGCAATWACGAGACAGTACAAAGGCGCAAGTTCA 923
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2285 tcaacttagcctgctctcaactcgtgctgctgctgctt 2327
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 924 TCACCTTCAGCATGCTGCTGCTACCTCATTTCTCGCTGCTCTT 966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: June 11, 2002, 05:49:17
Job time: 7093 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      June 11, 2002, 05:02:29 ; Search time 64.37 Seconds
```

9639.121 Million cell updates/sec

Title: US-09-819-946-1

Sequence: 1 atgctgtctctgcacgctcg.....ggcgcctgcgcgtccacctga 2526

Scoring table:	IDENTITY_NUC	Cancer 1 0
Cancer 10 0		

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 7670666

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
------------------	---------------	----

Listing first 45 summaries

```
Database : Issued_Patents_NA:*
```

```

1: /cgn2.6/ptodata/1/1na/5A_COMB.seq: *
2: /cgn2.6/ptodata/1/1na/5B_COMB.seq: *
3: /cgn2.6/ptodata/1/1na/6A_COMB.seq: *
4: /cgn2.6/ptodata/1/1na/6B_COMB.seq: *
5: /cgn2.6/ptodata/1/1na/PCRTUS_COMB.seq: *
6: /cgn2.6/ptodata/1/1na/backfiles1.seq: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	147.6	5.8	5275	1	US-08-485-588-1	Sequence 1, Appl1
2	147.6	5.8	5275	1	US-08-484-565-1	Sequence 1, Appl1
3	147.6	5.8	5275	2	US-08-480-751-1	Sequence 1, Appl1
4	147.6	5.8	5275	2	US-08-943-986-1	Sequence 1, Appl1
5	147.6	5.8	5275	3	US-08-353-784-1	Sequence 1, Appl1
6	147.6	5.8	5275	3	US-08-484-7198-1	Sequence 1, Appl1
7	147.6	5.8	5275	4	US-08-484-159-1	Sequence 1, Appl1
8	141.2	5.6	3809	1	US-08-485-588-3	Sequence 3, Appl1
9	141.2	5.6	3809	1	US-08-484-565-3	Sequence 3, Appl1
10	141.2	5.6	3809	2	US-08-480-751-3	Sequence 3, Appl1
11	141.2	5.6	3809	2	US-08-943-986-3	Sequence 3, Appl1
12	141.2	5.6	3809	3	US-08-353-784-3	Sequence 3, Appl1
13	141.2	5.6	3809	4	US-08-484-7198-3	Sequence 3, Appl1
14	141.2	5.6	3809	4	US-08-484-565-2	Sequence 2, Appl1
15	141.2	5.6	3809	4	US-08-484-159-3	Sequence 3, Appl1
16	133.8	5.3	4000	2	US-08-687-2894-2	Sequence 2, Appl1
17	133.8	5.3	5006	1	US-08-485-588-2	Sequence 2, Appl1
18	133.8	5.3	5006	1	US-08-484-565-2	Sequence 2, Appl1
19	133.8	5.3	5006	2	US-08-480-751-2	Sequence 2, Appl1
20	133.8	5.3	5006	2	US-08-943-986-2	Sequence 2, Appl1
21	133.8	5.3	5006	3	US-08-353-784-2	Sequence 2, Appl1
22	133.8	5.3	5006	3	US-08-484-7198-2	Sequence 2, Appl1
23	133.8	5.3	5006	4	US-08-546-998-1	Sequence 1, Appl1
24	133.8	5.3	5006	4	US-08-484-159-2	Sequence 2, Appl1
25	133.2	5.3	4134	4	US-09-162-0218-1	Sequence 1, Appl1
26	132.2	5.2	4131	1	US-08-485-588-4	Sequence 4, Appl1
27	132.2	5.2	4131	1	US-08-484-565-4	Sequence 4, Appl1

28	133.-2	5.2	4131	2	US-08-480-751.-4	Sequence 4, Appl1
29	133.-2	5.2	4131	2	US-08-443-986.-4	Sequence 4, Appl1
30	133.-2	5.2	4131	3	US-08-453-784.-4	Sequence 4, Appl1
31	133.-2	5.2	4131	3	US-08-484-719B.-4	Sequence 4, Appl1
32	133.-2	5.2	4131	4	US-08-484-159.-4	Sequence 4, Appl1
33	123	4.9	3177	4	US-09-134-513.-1	Sequence 1, Appl1
34	112.-2	4.4	3364	2	US-08-687-285A.-1	Sequence 1, Appl1
35	112	4.4	2148	5	PCT-US93-01642.-1	Sequence 1, Appl1
36	83.-2	3.3	2635	3	US-09-126-280.-3	Sequence 1, Appl1
37	83.-2	3.3	2670	3	US-09-126-280.-1	Sequence 1, Appl1
38	83.-2	3.3	2961	2	US-08-407-875.-1	Sequence 1, Appl1
39	78	3.1	3919	1	US-08-072-574.-5	Sequence 5, Appl1
40	78	3.1	3919	1	US-08-486-270.-5	Sequence 5, Appl1
41	78	3.1	3919	3	US-08-367-264.-5	Sequence 5, Appl1
42	77.-2	3.0	2736	4	US-08-617-785.-1	Sequence 1, Appl1
43	75.-6	3.0	4095	5	PCT-US91-09422.-18	Sequence 18, Appl1
44	75.-2	3.0	2619	3	US-08-537-797A.-1	Sequence 3, Appl1
45	75.-2	3.0	2619	3	US-08-537-797A.-3	Sequence 3, Appl1

ALIGNMENTS

RESULT
 1
 US-08-485-588-1
 Sequence 1, Application US/08485588
 Patent No. 5688938
 GENERAL INFORMATION:
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: Forrest H. Fuller
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 MOLECULES
 NUMBER OF INVENTION: 20
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,588
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 9
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827
 FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451

```

; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 955-0440
; TELEFAX: (213) 489-1600
; TELETYPE: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5275 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 515..3769
; OTHER INFORMATION:
;
US-08-485-588-1

```

```

Query Match      5.8%; Score 147.6; DB 1; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

```

```

QY 1426 ctaaacataatgaagcccaaatccagtgcaaggaagacacagtgcttaagtct 1485
DB 2078 cttcttatcatatgataaataattctgtgagtgatctcaaggagagcttcttc 2137
QY 1486 gtgtgtcagcagctgtcttgaaggacac---cagcagtggttgaagttccatcac 1542
DB 2138 AACTGAGTGAAGACACGCGCCGCGAGGACGAGAAAGATATCTAGGGGAGGCCACC 2197
QY 1543 tgcgtcttgagtggtgtgcccgtgagggttggaaccttcccaagaagtgaccttac 1602
DB 2198 TGCTGCTTGAAGTGTGTGAGATCTCTGATGGGAGTACAGCGACAGATGCAAGT 2257
QY 1603 agatgcagcctgtgaggaagagtgagcactgaggaagcagcactcttcccg 1662
DB 2258 GCCTGTGATAGTCCCTGATGACTTGTGATCAATGAGAACACACTTCTCATGCGC 2317
QY 1663 cgcactgtgtgttcttggtcttgctgagcacacctctggtgtgtctgagcagtaac 1722
DB 2318 AAGAGATGAGATTTGTCGTGTGAGACGACCTTGGGATGCGACTCAGCTTTGCT 2377
QY 1723 agcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1782
DB 2378 GTGCTGGGCAATTTTCTCAACAGCCTTCGCTGGGCTCTTCAATCAAGTCCGCAACAGC 2437
QY 1783 cctgtgtgagtgacagcagggcgccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1842
DB 2438 CCCATGCTCAAGGCCACCAACGCGAGCTCTCTATCTCTTCTTCTTCTTCTTCTG 2497
QY 1843 ggtagtgcagcctctatgtcttcttgggaaaccacaagcctgtgtgtgtgtgtgt 1902
DB 2498 TGCTTCTCAAGCTCCCTGTTCTTCAATCGGGAGCCCGACAGACATGAGAGTGGCCGCGC 2557
QY 1903 caggcccttcttgccttgggtttacacatcttctgtctgtctgtcagtgatctatc 1962
DB 2558 CACCCGCGCTTGGATCAAGCTTCGCTCTGATCTCGCATCTCTGGAAGAAACCAAT 2617
QY 1963 caactatcatcatcttcaagtttccaccaaggtacatcatcatcaccagcgtgtgtc 2022
DB 2618 CGGCTCTCTCTGTGTGTAG---GCCAAGATTTCCACACGCTTCCACCGGAAGTGTG 2674
QY 2023 caaacacacagtgctgagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2082
DB 2675 GGGCTCAACCTGAGTTCGCTGCTGCTTCTCTGACACCTTCAATGAGATTTGATCTGT 2734
QY 2083 ctactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2142

```

```

DB 2735 GCCATTTGGCTCAATACAGCGCCGCCCTGAGCTACCGCAACCGAGCTGAGACGAG 2794
QY 2143 ctgtgtatgcttgatgcacagagacacacccctgtgtgtgtgtgtgtgtgtgtgtgtgt 2202
DB 2795 ATCATCTTATCATACCTGCCACAGAGGCTGCTCATATGCGCTGGGCTTCTGTATGCGCTAC 2854
QY 2203 aatgacctctctcactcagcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
DB 2855 ACCGTGCTGCTGGCGCCCATATCTCTTCTTCTGCTTCAAGTCCGGAAGCTGCGCAGAG 2914
QY 2263 aactaacagagagcacaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2322
DB 2915 AACTTCAATGAACCAAGTTCAATCACTTCAATCATCTCTCTCTCTCTCTCTCTCTCT 2974
QY 2323 gctcttaccacagcagcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2382
DB 2975 TCTTTATCTCCCGCTTACCGCAGACTTACGCGCAAGTGTCTCTCTCTCTCTCTCTCTCT 3034
QY 2383 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2442
DB 3035 GCCATCTGGCGGCGCAGCTTTGGCTGTGCGCTGTATCTCTTCAACAAGTCTGATC 3094
QY 2443 atcctctgcgcagcagcactcaacagcagcagcagcagcagcagcagcagcagcagc 2482
DB 3095 ATCTCTTCAAGCTTCCCGAAGACCATCAGAGAGTGC 3134

```

RESULT 2

```

US-08-484-565-1
; Sequence 1, Application US/08484565
; Patent No. 5763569

```

```

; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCULON RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/484,565
; FILING DATE: 7 June, 1995

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 9

```

```

; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994

```

```

; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994

```

```

; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994

```

```

; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993

```

```

; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993

```

```

; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993

```

```

; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992

```


FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Heber, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 213/004
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ. ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5275 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 515..3769
 OTHER INFORMATION:
 US-08-480-751-1

Query Match 5.8%; Score 147.6; DB 2; Length 5275;
 Best Local Similarity 47.5%; Pred. No. 1.7e-28;
 Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

QY 1426 ctaaacataatgaagacaaatccagtgagcaaggaaggaacacagtgagcttaagtc 1485
 DB 2078 ctcttatcatatgataaaaaattctgtgagatgctcaaggagagctcttctcc 2137
 QY 1486 gtgtgtccagcgacgtcttgaaggacac---cagcgaagtggttaacggttccatca 1542
 DB 2138 AACTGCAGTCGACACTGCTCGGAGGAGACAGAAAGATCTATTGAGGGGAGCCACAC 2197
 QY 1543 tgcgtcttgagtggtgtgcccgtggtgagcttccctcaagaagagtagccttacc 1602
 DB 2198 tgcgtcttgagtggtgtgagagtgctgagtgagtgagtgagtgagtgagtgag 2257
 QY 1603 agtgcagagcgtgtggaagaaagagtgagcaacctgaaggaagcgaagacctctccg 1662
 DB 2258 GCCTGATATAGTGCCTGTGACTTCTGTGATGATGATGATGATGATGATGATGATG 2317
 QY 1663 cgcactgt 1722
 DB 2318 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2377
 QY 1723 agcgtctgt 1782
 DB 2378 GTGCTGGGATTTTCTTCACAGCCTTCGTGCTGGGCTTCATCAAGTTCGCCACACG 2437
 QY 1783 cctgt 1842
 DB 2438 CCCATGTCAGGCAACCAACCGGAGCTCTCTATCTCTTCTTCTTCTTCTTCTTCTG 2497
 QY 1843 ggtagt 1902
 DB 2498 TGTCTTCAGCTCCCTGCTTCATGCGGAGCCACAGACATGAGACGTCCGCTGCGC 2557
 QY 1903 cagggcctctgtccctgtgttcaacatcttctgtctgtcgtgaagtgctcaatc 1962
 DB 2558 CACCCGCGCTTTGGCATACAGCTTCGTGCTCTGATCTCGGATCTCGTGGAACCAAT 2617
 QY 1963 caactatatacatcttcaagtttccacaaggtactatacttaccacgctgtgtc 2022

DB 2618 CGGTCCTCCTGGTGTGTGAG---GCCAAGATCCCAACCAAGCTTCCACCGGAAGTGAG 2674
 QY 2023 caaacacagtgctgtgcccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2082
 DB 2675 GGGCTAACCTGCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2734
 QY 2083 ctactgt 2142
 DB 2735 GCCATTGGCTCAATACAGGCGCCCTCGAGCTACCGCAACACAGCTGAGAGAGAG 2794
 QY 2143 ctgtgtatgt 2202
 DB 2795 ATCATCTTATCATCCCTACGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2854
 QY 2203 aatgctctctctcaatgaagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
 DB 2855 ACCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2914
 QY 2263 aactaacagagcacaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2322
 DB 2915 AACTCAATGAAGCAAGTTCATCAGCTTCAGCATCTCATCTTCTTCAATCGTGTGATC 2974
 QY 2323 gcttcttcaacacagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2382
 DB 2975 TCTTCAATCCCGCTACGACGACACTTACGCAAGTTCGCTGCTGCTGCTGCTGCTG 3034
 QY 2383 gctgagcgtgacagcgt 2442
 DB 3035 GCCATCTGCGGCGGCAAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3094
 QY 2443 atcctctgcccagcagcctcaacagacagagagacacctcc 2482
 DB 3095 ATCCTCTTCAAGCCTTCGCGAACCACATCAGAGAGTGC 3134

RESULT 4
 US-08-943-986-1
 ; Sequence 1, Application US/08943986
 ; Patent No. 5962314
 ; GENERAL INFORMATION:
 ; APPLICANT: Edward M. Brown
 ; APPLICANT: Steven C. Hebert
 ; APPLICANT: James E. Garrett, Jr.
 ; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: First Interstate World Center
 ; STREET: Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: FASTSEQ
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/943,986
 ; FILING DATE: 03-OCT-1997
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/484,565
 ; FILING DATE: 7-June-1995
 ; APPLICATION NUMBER: 08/353,784
 ; FILING DATE: 9 December, 1994
 ; APPLICATION NUMBER: PCT/US/94/12117
 ; FILING DATE: 21 October, 1994
 ; APPLICATION NUMBER: U.S. 08/292,827
 ; FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-943-986-1

Query Match 5.8%, Score 147.6; DB 2; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

QY 1426 ctaaacataatgagacaaatcagtgagcaggaagaagacacagtgctcctaaagtct 1485
DB 2078 CTCTTCATCAATGATGAAAAATTTCTGTGAGTGTGATTCCTCAAGGAGGTGCTTTCCTC 2137
QY 1486 gtgtgtccagcagactgtcttgaagagcagc---cagcagtggttaagtggttcataac 1542
DB 2138 AACTGACATCGAGACTGCTGTGAGGAGACACAGAAAGCAATATTGAGAGGGAGCCACACC 2197
QY 1543 tgcgtcttgagtggtgtgcccgtgagggcctggaccttccctcaagaagtgacctctac 1602
DB 2198 TGCTGCTTGAAGTGTGTGAATGTCTGATGGGAGTACAGCGACGAGACAGATGCAAGT 2257
QY 1603 agatgccagcctgtgtggaagaagaagtgggacactgaagggagacagacctgtctccg 1662
DB 2258 GCGCTGTGATGAGTGCCTGATGACTTCTGTCATGAGAACCAACACTTCCGTCATCGCC 2317
QY 1663 cgcacgtgtgtgttcttgcttgctgagcaacacctctgtgtgtctgtgcaagctaac 1722
DB 2318 AAGGAGATCGAGTTCCTGTCTCGGAGACGAGCCCTTCGGGATCGCATCGCTCTTTGCT 2377
QY 1723 acgctgtcgtcgtcgtcgtcgttggaactgtcgtcgtgttgcctgtgcaactagacac 1782
DB 2378 GTGCTGGGCAATTTCTCATGACAGCCCTTGTGCTGGCGCTTCCTCATCAAGTTCCGCAACG 2437
QY 1783 cctgtgtgtgagtgagcagagggggccgctgtgtcttctatgtcgtgtgtcctctgcaaga 1842
DB 2438 CCCATCGTCAGGCGACCAACCGGAGCTCTTCATCTCTCTCTCTCTCTCTCTCTCTCTGCTGC 2497
QY 1843 ggtatgtgagcagcctcctccttctgtggaaccacacaagccgcgtcgtcgtctagcgc 1902
DB 2498 TGCTTCTCCAGTCTGCGC 2557
QY 1903 cagcgcctccttgcctcctgttctacacatcctcctcctcctcgtcgtcgtcgtcgtcgtc 1962
DB 2558 CAGCGCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2617

QY 1963 caactaatcatcctcctcaagtttccaccaaggtaccctacatcttaccacgcctgggtc 2022
DB 2618 CGGGTCTCTCTGCTGTTTGA---GCCAAGATTCCACACAGCTTCCACCGGAGAGTG 2674
QY 2023 caaaacacagtgctgtcgtcgtgtgtgtatcagctcagcggccagcgtgtctatcgt 2082
DB 2675 GGGCTCAACCTGACAGTCTCTGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2734
QY 2083 ctaacttgctgt 2142
DB 2735 GCCATTTGGCTCAATACAGCGCCCTCTGAGCTACCGCAACAGAGAGCTGGAGAGAG 2794
QY 2143 ctgtgtatgt 2202
DB 2795 ATCATCTTCATACCTGCGACAGAGGCTGCTCATATGCGCTGCTCTCTGATTCGGCTAC 2854
QY 2203 aatgacctctctcctcctcctcctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2262
DB 2855 ACCGTGCTGCTGCGCGCATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2914
QY 2263 aactacaagagggccaatgtgtacacctcagcctcctcctcctcctcctcctcctcctcctc 2322
DB 2915 AACTTCATGAGGCGCAAGTTCATCACTTACATGCTCTCTCTCTCTCTCTCTCTCTCTCTGATC 2974
QY 2323 gctctctcaccagcgagcgtctcagcagcagatcctcctcgtcgtcgtcgtcgtcgtcgtc 2382
DB 2975 TCTTCATCTCCCGCTACCGCCACACTTACGCAAGTTCGTCTGCTCTGAGGATGATC 3034
QY 2383 gctggtcgtgagcagcctgagcagcgtcgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2442
DB 3035 GCCATCTGCGCGCGCATCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3094
QY 2443 atcctctgcgcgccagacctcaacagcacagacacctcc 2482
DB 3095 ATCTCTTCAACCTTCCCGGAACACATGAGGAGGTGC 3134

RESULT 5
US-08-353-784-1
Sequence 1, Application US/08353784
Patent No. 6011068
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert, Manuel
APPLICANT: Bradford C. Van Wageningen, Manuel
APPLICANT: F. Balandieri, Forrest H. Fuller,
APPLICANT: Eric G. Delmar, and Scott T. Moe
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,784
FILING DATE: 9 December, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below: 8
APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 209/069
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-353-784-1

Query Match 5.8%; Score 147.6; DB 3; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;

QY 1426 ctaaaataatgagacccaataaccagtgcaaggaaggaacacagtgcttaagtc 1485
DB 2078 CTCCTTATCAATGATGATAAAATTCGTGAGTTCATCAAGGAGGCTTTCTCC 2137
QY 1486 ggtgttcacagcagctgtctgaagggcac---cagcaggtgttaccggtttccatcac 1542
DB 2138 AACTGCAGTCGAGACTGCTCTGGCAGGAGGACCAGAAAGGATCTTGAAGGGGAGCCAC 2197
QY 1543 tgcgtctgaagtgtgtccctgtgaggtcgtggaaccttcctcaacaagaagtgaacctctac 1602
DB 2198 TGCTGCTTGAAGTGTGAGATGCTCTGATGGGAGTACAGCAGACAGATGCAAGT 2257
QY 1603 agatgcagacctgtggaagaagaagtgtggcaactgtgaagggaagcaagacctgtctccg 1662
DB 2258 GCGTGTGATGAAGGCCCTGATGCTGTGTCATGAAGCAACCACTTCTGCATCGCC 2317
QY 1663 cgcacgtgtgtgtttgtgcttgcgtgagcagacctcttgggtgtgtcgtgagctaac 1722
DB 2318 AAGGAATGAGATTTCTGTCGTGAGGAGCCCTTGGGATCGCATCGACGCTTTTGTCT 2377
QY 1723 acgtgtcgtcgtcgtcgtcgtcgtggaactgtgacctgtgttgcctgagcactagacac 1782
DB 2378 GTGCTGGGATTTTCTCAACAGCCTTCGTGGGCTGTCTATCAAGTTCCGCAACAG 2437
QY 1783 cctgtgtgaagtcaagcagggggcgccgtgtgtcttctatagctgaggtccctgagca 1842
DB 2438 CCATGTGTAAGGCCCAACCGGAGCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCT 2497
QY 1843 ggtagtgagcagacctatagctcttcttgggaacccacaagacctgtggtctgagc 1902
DB 2498 TGCTTCTCAAGCTCCTGTCTTCTCAATCGGGAGCCCGCAGACTGAGCTGCGCTGCGC 2557

QY 1903 cagggcctcttggccttgggtttccacatcttccttgcctgctgacagttcgtctatcc 1962
DB 2558 CAGCGGCTTTGGCATGACGCTTCGTCGTGATGCTGATGCTGATGCTGATGCTGATGCT 2617
QY 1963 caactaatcatcatcttcaagtttccaccaaggttaactatcttaccagcctgggtc 2022
DB 2618 CGGCTCTCTCTGTTGAG---GCCAAGATTCCACACAGCTTCCACCGAAGTGGTG 2674
QY 2023 caaaccagcagcgtcgtgtgttggatgacagctcagcggccagctgtctatctgt 2082
DB 2675 GGGCTAACCTGACGATTCCTGCTGCTGCTCTCTGACCTTCAAGTATGAGTTGTCTAT 2734
QY 2083 ctactgtgctgt 2142
DB 2735 GCATTTGGCTCAATACAGCGGCCCTCGAGCTACCGAACACAGAGCTGGAGAGCAG 2794
QY 2143 ctgtgtatgctgt 2202
DB 2795 ATCATCTTATCACTCCGACGAGGAGGCTGCTCATGGCGCTGCTGCTGCTGCTGCT 2854
QY 2203 aatgacctctctccatcagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
DB 2855 ACCTGCTTGTGGCGGCGCATGCTGCTTCTTCTGCTTCAAGTCCGAGAGCTGCGAGAG 2914
QY 2263 aactaacagcagcacaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2322
DB 2915 AACTCAATGAAGCAAGTTCATCACTTACAGATCTCATCTTCTTCAATGCTGTGATC 2974
QY 2323 gctcttccacagcagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2382
DB 2975 TCTTTATCCCGCTACGCGCAGCACTTACGCAATGCTGCTGCTGCTGCTGCTGCTG 3034
QY 2383 gctgagctgtgagcagcctgtgagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2442
DB 3035 GGCATCTGTGGCGGCGCATGCTTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3094
QY 2443 atctctgcgcgcagacctcaacagcagcagcagcagcagcagcagcagcagcagcag 2482
DB 3095 ATCTCTTCAAGCTTCCCGGAACACCATGAGAGAGTGC 3134

RESULT 6
US-08-484-719B-1
; Sequence 1, Application US/08484719B
; Patent No. 6031003
; GENERAL INFORMATION:
; APPLICANT: Edward F. Nemeth, Edward M.
; APPLICANT: Brown, Steven C. Hebert,
; APPLICANT: Bradford C. Van Wageningen,
; APPLICANT: Manuel F. Baladrin,
; APPLICANT: Forrest H. Fuller, Eric G.
; APPLICANT: Delmar, Scott T. Moe
; TITLE OF INVENTION: CALCIUM RECEPTOR ACTIVE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; SUITE: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS Word
; SOFTWARE: FASTSEQ for Windows Version 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,719B
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Douglas C. Murdock
REGISTRATION NUMBER: 37,549
REFERENCE/DOCKET NUMBER: 213/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
US-08-484-719B-1

Query Match 5.8%; Score 147.6; DB 3; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;
Qy 1426 ctaaacataatgagacccaataatccagctgagcaggaagaacaccaggtgcctaagctt 1485
Db 2078 CTCTTCATCAATGAGAAATTTCTGTGAGCTGATTCTCAAGGAGGTGCTTTCTCC 2137
Qy 1486 gtgtgtccagcagactgtcttgaaggacac--cagcgagtgttaacgggttccatcac 1542
Db 2138 AACTGCATCCGAGACTGCTGCGACGAGGACAGAAAGCATATTGAGGGGAGGCCACC 2197
Qy 1543 tgcgtcttgagtggtgcccctgtggggtggtgaccttccctcaagaagtgaaccttacc 1602
Db 2198 TGCTGCTTTGAGTGTGTGAATGTCTGATGGGAGTACAGCGACGACAGATCAAGT 2257
Qy 1603 agatcccgccctgtgggaagaagaagtgggacctgtggggaagcagaccgtctccg 1662
Db 2258 GCCTGTGTAAGTGCCTGATGACTTCTGTGTCCAATGAGAACCACTTCCTGCAATCGC 2317
Qy 1663 cgaactgtgtgttttctgtgtgtgtgacacacctcttgggtgtgtgtgacagctaac 1722
Db 2318 AAGGAGATCGAGTTCCTGTCGTGACCGAGCCCTTCGGATGCACTACGCTCTTTGCT 2377
Qy 1723 acgctgtgctgtgctgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1782
Db 2378 GTGCTGGGCAATTTCTCACAAGCCTTCGTGCTGGCGCTTCATCAAGTTCCGCAACAG 2437
Qy 1783 cctgt 1842
Db 2438 CCATGCTCAAGGCGCAACCGGAGGCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTGC 2497
Qy 1843 ggtagtgcagcctctatgtcttcttggggaaccacaaagcctgtgtgtgtgtgtgtgc 1902

Db 2498 TCTTCTCCAGCTCCCTGTTCTTCAACGGGAGGCCACAGACTGACGTCGCGCTGGCC 2557
Qy 1903 cagggcccttggccctgtgtttcaacatcttcctgtctgtcctgaagttgtctatc 1962
Db 2558 CAGCGGCTTGGGATCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2617
Qy 1963 caactatcatatcttcaagtttccaccaggtgaactcaacttaccagcctgtgtgc 2022
Db 2618 CGGCTCTCTGCTGTTTGAAG--GCCAAGATTCCACACGCTTCCACCGGAAGTGTG 2674
Qy 2023 caaacaccagtgctgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2082
Db 2675 GGGCTCAACCTGACGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2734
Qy 2083 ctaactgt 2142
Db 2735 GCCATTTGGCTCAATACAGCGCCCTGAGCTACCCGACACGACGAGCTGGAGCAG 2794
Qy 2143 ctgt 2202
Db 2795 ATCATCTTATACCTGCGCACAGGGCTGCTCATGGCGCTGGGCTTCTGATCGGCTAC 2854
Qy 2203 aatggcctcctccatcagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
Db 2855 ACCTGCTTCTGCGCGCCCATCTGCTTCTTCTGCTTCAAGTCCCGAAGCTGCCAGAG 2914
Qy 2263 aactacaagagccaatgyltcaacttcaagctgtcttcaacttctgtgtgtgtgt 2322
Db 2915 AACTTCATGAGCAAGTTACCTTACACATCTCATCTCTCTCTCTCTCTCTCTCTCT 2974
Qy 2323 gcttcttaccagggcaggtgtcagcgaggaagttactgtcctgtgtgtgtgtgtgt 2382
Db 2975 TCTTTCATCCCGCCCTACCGCACCTTACGCAAGTGTCTTCTGCTGTGAGGCTGATC 3034
Qy 2383 gctggctgtgagcagcctgtgagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2442
Db 3035 GCCATCCTGGGCGGCGGCGCTTGTGCTTGTGCTGCTGTGCTGTGCTGTGCTGTG 3094
Qy 2443 atcctctgcggccagacctcaacagcagcagcagcagcagcagcagcagcagcagc 2482
Db 3095 ATCTCTTCAAGCTTCCCGGAACACCATCGAGGAGGTGC 3134
RESULT 7
US-08-484-159-1
Sequence 1, Application US/08484159
Patent No. 6313146
GENERAL INFORMATION:
APPLICANT: Bradford C. Van Wageningen
APPLICANT: Manuel F. Balandrin
APPLICANT: Eric G. Del Mar
APPLICANT: Edward F. Nemeeth
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,159
FILING DATE: 7 June, 1995

```
CLASSIFICATION: 435
PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
APPLICATION NUMBER: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
APPLICATION DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
APPLICATION DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
APPLICATION DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
APPLICATION DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
APPLICATION DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
APPLICATION DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
APPLICATION DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
APPLICATION DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 214/101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-484-159-1

Query Match          5.8%; Score 147.6; DB 4; Length 5275;
Best Local Similarity 47.5%; Pred. No. 1.7e-28;
Matches 503; Conservative 1; Mismatches 550; Indels 6; Gaps 2;
```

```

Db 2438 CCCATCTCAAGGCGACCAACCGGAGCTCTCTATCTCTTCTCTCCCTGCTGC 2497
Qy 1843 gtagtgcagccctctatggtctctttgggaagcccaagagcctgtgtctgacgc 1902
Db 2498 TGCTTCTCCAGCTCCCTGTTCTTCATGCGGGAGCCCGAGACTGAGCTGCCCTCGC 2557
Qy 1903 cagccctcttgccttggtttccacctcttcctgcctgcctgacgttcctac 1962
Db 2558 CAGCGGCGCTTGCGATCAGCTTCGTCTGATCTCGATCTCGATCTCGATCTCGAT 2617
Qy 1963 caactatctatcttcaagtttccaccaggtactactatcttccacgcctggtc 2022
Db 2618 CGGCTCTCTGCTGGTTGAG--GCCAATATCCACACGCTTCCACCGGAAGTGGTG 2674
Qy 2023 caaacacagtgctgcgcctgtgtatgatcagctcagcgccacgccttatctg 2082
Db 2675 GGGCTCAACCTGCAAGTTCCTGCTGCTGCTTCTCTGCACTTCAAGCATGATGCTGT 2734
Qy 2083 ctactgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2142
Db 2735 GCCATTTGGCTCAATACAGCGCCCTCGAGCTACCGCAACACAGCTGAGAGACGAG 2794
Qy 2143 ctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2202
Db 2795 ATCATCTTCAATCAACCTGCAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2854
Qy 2203 aatgcctctctccatcagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
Db 2855 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2914
Qy 2263 aactacaagagccaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2322
Db 2915 AACTCAATGAAGCAAGTTCATCAGCTTCAAGTTCATCAGTTCATCAGTTCATCAG 2974
Qy 2323 gcttcttcaacagccagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2382
Db 2975 TCTTTCATCCCGCTCAAGGCGACACTTACGCAAGTTCGCTGCTGCTGCTGCTGCT 3034
Qy 2383 gctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2442
Db 3035 GCCATCTGCGGCGGCGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3094
Qy 2443 atcctctgcgcgcagcctcaacagcagcagcagcagcagcagcagcagcagc 2482
Db 3095 ATCCTCTTCAAGCCTTCCGGAACACCATGAGAGAGTGC 3134

RESULT 8
US-08-485-588-3
; Sequence 3, Application US/08485588
; Patent No. 5688938
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: Forrest H. Fuller
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-485-588-3

Query Match 5.6%; Score 141.2; DB 1; Length 3809;
Best Local Similarity 47.1%; Pred. No. 6.4e-27;
Matches 499; Conservative 1; Mismatches 554; Indels 6; Gaps 2;

QY 1426 ctaaacatgaagaccacaaatccagtgagcaagaaagcaaacaggtgcctaagtct 1485
DB 1933 CTCTTCATCAACGAGGAGAAATCTGTGAGTGTCTCCAGAGGAGTCCCTTCTCC 1992
QY 1486 gtgtgtccagcgagctgtcttgaaagggcaccagcgag---tggtgaagggttccatcac 1542
DB 1993 AACTGCAGCCGAGACTGCTGGCAGGACCGAGAAAGGATCATTTGAGAGGAGCCACC 2052
QY 1543 tgcgtcttgagtgctgtgcctgtgagggctggaacctcccaacaagagtgaccttac 1602
DB 2053 TGCAGCTTTGAGTGTGAGATGTCTGATGGGAGTATGATGAGACAGATGCCAGT 2112
QY 1603 agatgcagcgctgtggaagaagagtggaacctgagggagcagcgactgtcccg 1662
DB 2113 GCCGTGAACAAGTCCAGAGATCTCTGTGCAATGAGAACACACACTCTCGATTGCC 2172
QY 1663 cgacgtgtgtgtttgt 1722
DB 2173 AAGAGATCGAGTTTCTGTCTGTGACGAGCCCTTTGGAGTCGATCCCTCTTTGCC 2232
QY 1723 acgt 1782

DB 2233 GTGCTGGGCAATTTTCTTACACAGCCCTTGTGCTGGGTGTGTTATCAATTCGCAACACA 2292
QY 1783 cctgt 1842
DB 2293 CCAATTCGAGGAGCCACCAACAGAGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 2352
QY 1843 ggt 1902
DB 2353 TGCCTTCAGCTCCCTTCTTCATCGGGAGGCCCGAGACTGAGCTGCGCGCTGCGC 2412
QY 1903 caggccctcttggccctgt 1962
DB 2413 CAGCGCGCTTTGGATCAGCTGTGCTGTGATCATGATGATGATGATGATGATGATGAT 2472
QY 1963 caactaatcatcatccttaagtlttcgaacaaagtaaccttaaccttaacctgtgtc 2022
DB 2473 CGTGTCCCTCTGT 2529
QY 2023 caaaaccaggt 2082
DB 2530 GGGCTCAACCTGCAAGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2589
QY 2083 ctaactgt 2142
DB 2590 GTGATGTGCTCTACACCGCGCGCCCTCAAGCTACCGCAACAGAGACTGGAGATGAG 2649
QY 2143 ctgt 2202
DB 2650 ATCATCTTCATCAGCTGTGCAAGGCTCCCTCAAGGCTCGGGCTTCCTGATCGCTAC 2709
QY 2203 aatgagctcctctcatcagtgagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
DB 2710 AACTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2769
QY 2263 aactcaacagagccaaatgt 2322
DB 2770 AACTTCATGAGGAGCAAGTATCATCTTCAGATGCTCATCTTCATCTTCATCTTCAT 2829
QY 2323 gctctcttcaacagagcgagtgctgaagagtgagtgagtgagtgagtgagtgagtg 2382
DB 2830 TCCCTCATTCAGACCTATGACAGCACTATGCAAGTGTGTGTGTGTGTGTGTGTGT 2889
QY 2383 gctgagctgagcagctgtgagcagcgctgtgtgtgtgtgtgtgtgtgtgtgtgt 2442
DB 2890 GCCATCTGAGGAGCCAGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2949
QY 2443 atcctctgcgcccagacctcaacagcacagacacttc 2482
DB 2950 ATTCTCTCAAGCATTCGCCGAAACACATCGAGAGGTGC 2989

RESULT 9
US-08-484-565-3
Sequence 3, Application US/08484565
Patent No. 5763569
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:

```

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-484-565-3

```

```

Query Match          5.6%; Score 141.2; DB 1; Length 3809;
Best Local Similarity 47.1%; Pred. No. 6.4e-27;
Matches 499; Conservative 1; Mismatches 554; Indels 6; Gaps 2;

```

```

OY 1426 ctaacataatgagacccaatccagtgagcaaggaaggaacaccagtgccctaagtct 1485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1933 CTCTTCATCAACGAGAGAGAAATCCTGTGAGTGGCTTCACAGGAGAGTGCCTTCTC 1992
OY 1486 gtgtctcagcagctgtcttgaaggagcaccagcgag---tggttaaggtttccatc 1542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1993 AACTGCAGCGGAGACTGCTGCGAGGACCAAGAAAGGATCATTTGAGGGAGGCCACC 2052
OY 1543 tgcctcttgagtggtgcccctgtgagccttcccaacaagagtgacctctac 1602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2053 TGCTGCTTGAATGTGTGAGTGTCTCTGATGGGAGTATAGTATAGACATGCGCACT 2112
OY 1603 agatgccagccttggtggaagaagatggtgcaacctgaagggaagcaccgtgtctccg 1662
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2113 GCCTGTAAACAGTCCCAATGACTTGTGTCGAATGAGAACACACCTCTGCAATTGCC 2172
OY 1663 gcgactgtgtgttttggcttggtgagcacctctgtgtgtgtgtgtgtgtgtgtgtgt 1722
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 2173 AAGGAGATCGAGTTTCTGTCTGCGAGGAGGCCCTTTGGATCGCACTACCTCTTTGCC 2232
OY 1723 acgctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 1782
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2233 GTGCTGGGCAATTTCTCTGACAGCCTTTGTGCTGAGTGTGTTTATCAAGTTCCGCAACA 2292
OY 1783 cctgtgtgaagtcaagcaaggagcgccgtgtcttctctatgtctgtgtctgtctgtca 1842
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2293 CCCATTGTCAAGGACCAACCAACGAGAGCTCTCTACCTCTCTCTCTCTCTCTCTCTCT 2352
OY 1843 ggtgaagcagcctctatgtctcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2353 TGCTTCTCAAGCTCCCTGTTCATCGGGAGGCCAGGAGCTGAGCTGCCCTGCCCTGCC 2412
OY 1903 caagccctctgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1962
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2413 CAGCCGCGCTTTGGCATACAGCTTCGTCCTGTGATTCATGCTCTGTGTGTAACACAC 2472
OY 1963 caactaatacatcttcaagtltccaccaaggtactcaactcaactcaactcaactcaact 2022
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2473 CGTGTCTCTCTGCTGTTTGTAGG---CCAAGATCCCAACAGCTTCCACCGCAAGTGCTG 2529
OY 2023 caaacaacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2082
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2530 GGGCTCAACCTGCAGTTCCTGCTGTTTCTCTGACCTTCATGAGATGTCTATCTGT 2589
OY 2083 ctactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2590 GTGATGTGGCTTACACCGCGGCCCTCTCAAGCTACCGCAACGAGAGCTGAGATGTAG 2649
OY 2143 ctgtgtatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2650 ATCATCTTCATACGTGCGACGAGGAGGCTCCCTCATGTGCGCTGTGATGCGCTAC 2709
OY 2203 aatgtctctctctcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2710 ACCTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2769
OY 2263 aactacaagagcaacaatggtgtcaactctcaactctcaactctcaactctcaactctca 2322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2770 AACTTCAATGAGCGCAAGTTTCATACCTCTCAGATCTCATCTCTTCAATCGTGTGATC 2829
OY 2323 gcccttcaacagcgagcggtctcaagaggaagtaactctgtctgtgtgtgtgtgtgtgt 2382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2830 TCTTTCATTCACGACTATGCGACGACCTATGCAAGTTTGTCTGCTGCTGAGGTGAT 2889
OY 2383 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2890 GCCATCTGTGACGCGACGCTTTGGCTTGTGCGTGTGATCTTTCACACAGATCTACATC 2949
OY 2443 atccctgtccgcccagacctcaacagacagagcaactctc 2482
DB 2950 ATTCTTCAAGCATTCCGCAACACCATTCAGAGAGGTGC 2989

```

```

RESULT 10
US-08-480-751-3
; Sequence 3, Application us/08480751
; Patent No. 5858684
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street

```


APPLICANT: Bradford C. Van Wagenen,
APPLICANT: Manuel F. Balandrin,
APPLICANT: Forrest H. Fuller, Eric G.
APPLICANT: Delmar, Scott T. Moe
TITLE OF INVENTION: CALCIUM RECEPTOR ACTIVE
MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS Word
SOFTWARE: FASTSEQ for Windows Version 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,719B
FILING DATE: 7 June, 1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Douglas C. Murdock
REGISTRATION NUMBER: 37,549
REFERENCE/DOCKET NUMBER: 213/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373...3606
US-08-484-719B-3

Query Match 5.6%; Score 141.2; DB 3; Length 3809;
Best Local Similarity 47.1%; Pred. No. 6,4e-27;
Matches 499; Conservative 1; Mismatches 554; Indels 6; Gaps 2;

QY 1426 ctaacataaatgatgacccaatccagtgccagcaaggaagacacacagtgcttaagtc 1485
DB 1933 CTCCTCATCAACGAGAGAAATCTCTGTGAGTGCTTCTCCAGGAGGTGCCCTTCTCC 1992

QY 1486 gtagtccagcagcagctgtcttgaaggccacgagag---tggtacgggttccatcac 1542
DB 1993 AACCTGACGCCAGACTGCTGGCAGGACGAGAAAGGATCATTTGAGGGGAGGCCACC 2052
QY 1543 tgcgtttgagtggtgtgtccctgtggggttggaaccttccataaagtgtaacctcac 1602
DB 2053 TGCCTGTTGAGTGAGTGAGAGTCTCTGATGGGAGTATAGTATGACAGATGCCAAT 2112
QY 1603 agatccagcctgttggaagaagtggtgacccggaaggaagcagacgtctcccg 1662
DB 2113 GCCTGTACAGAGGCCAGATGACTTGTGTCATGAGAACACACACCTCTCATTTGGC 2172
QY 1663 cgcactgtgtgttttggccttgctgagacacaccttgggtgctgctgagactaac 1722
DB 2173 AAGGATGAGATGCTGTGCTGTGAGAGGAGCCCTTGGGATGCACTCACCTCTTGGC 2232
QY 1723 agcgtctgctgt 1782
DB 2233 GTGCTGGCATTTTCTTGACAGCCTTTGTGCTGGGTGTGTGTGTGTGTGTGTGTGTGT 2292
QY 1783 cctgtgtgagtgcaagcaggggagcgcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1842
DB 2293 CCATTTGTCAAGGCCACCAACGAGAGCTCTCTACTCTCTCTCTCTCTCTCTCTCTCTCT 2352
QY 1843 gtagtgagcagcctctatgt 1902
DB 2353 TGCTTCTCAGCTCCCTGTCTTCATCGGGAGGCCACGAGTGAAGTGGCCGCTGCGC 2412
QY 1903 cagagcctctgt 1962
DB 2413 CAGCCGCTTTGTGATCAGCTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2472
QY 1963 caactaatcatcatcctcaagtttccacaaggtactcaacttaccacagcgtgtgtgt 2022
DB 2473 CGTGTCTCTCTGT 2529
QY 2023 caaaacacagtggt 2082
DB 2530 GGGCTTCAACCTGAGT 2589
QY 2083 ctactgt 2142
DB 2590 GTGATGTGCTGTACACCGCCGCCCTCAAGTACCAGACAGAGAGCTGAGATGTAG 2649
QY 2143 ctgt 2202
DB 2650 ATGATCTTATCAGGTGACGACGAGGAGGCTCCATGATGAGGCTGTGATGCGCTAC 2709
QY 2203 aatggtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2262
DB 2710 ACCGTGCTGT 2769
QY 2263 aactacaagagcgaatggt 2322
DB 2770 AACCTTAATGAAAGCAAGTTTCACTTCAAGATCTATCTTCTTCACTGCTGTGATC 2829
QY 2323 gcttcttcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2382
DB 2830 TCTTTCATTCAGCCATATGACGACACCTATGGAAGTTGTCTGCTCCGTAAGAGTGAT 2889
QY 2383 gctgt 2442
DB 2890 GCCATCTGTGAGCCAGCTTTGCTGTGAGGCTGTGATCTTCTTCAACAAGATCTATC 2949
QY 2443 atcctctgcccagaccttaacacagcagcagcagcagcagcagcagcagcagcagcagc 2482
DB 2950 ATTCTTCAAGCCATCCGCAACACATCGAGAGGTGC 2989

RESULT 14
US-08-546-998-2
; Sequence 2, Application US/08546998
; Patent No. 6211244


```

1 GENERAL INFORMATION:
2 APPLICANT: Van Wageningen, Bradford C.
3 APPLICANT: Moe, Scott T.
4 APPLICANT: Balandrin, Manuel F.
5 APPLICANT: Delmar, Eric G.
6 APPLICANT: Nemeth, Edward F.
7 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
8 COMPOUNDS
9 NUMBER OF SEQUENCES: 2
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Lyon & Lyon
12 STREET: First Interstate World Center
13 STREET: Suite 4700
14 STREET: 633 West Fifth Street
15 CITY: Los Angeles
16 STATE: California
17 COUNTRY: USA
18 ZIP: 90071
19
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
22 COMPUTER: IBM PC compatible
23 OPERATING SYSTEM: PC-DOS/MS-DOS
24 SOFTWARE: FASTSEQ
25
26 CURRENT APPLICATION DATA:
27 APPLICATION NUMBER: US/08/546,998
28 FILING DATE: October 23, 1995
29
30 CLASSIFICATION: 424
31 PRIOR APPLICATION DATA: including application
32 PRIOR APPLICATION DATA: described below:
33 APPLICATION NUMBER:
34 FILING DATE:
35
36 ATTORNEY/AGENT INFORMATION:
37 NAME: Heber, Sheldon O.
38 REGISTRATION NUMBER: 38,179
39 REFERENCE/DOCKET NUMBER: 215/304
40 TELECOMMUNICATION INFORMATION:
41 TELEPHONE: (213) 489-1600
42 TELEFAX: (213) 955-0440
43
44 TELEX: 67-3510
45
46 INFORMATION FOR SEQ ID NO: 2:
47
48 SEQUENCE CHARACTERISTICS:
49 LENGTH: 3809 base pairs
50 TYPE: nucleic acid
51 STRANDEDNESS: single
52 TOPOLOGY: linear
53
54 MOLECULE TYPE: cDNA to mRNA
55
56 FEATURE:
57 NAME/KEY: CDS
58 LOCATION: 373...3606
59 OTHER INFORMATION:
60
61 US-08-546-998-2
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945

```

[illegible]

STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,159
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 214/101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-484-159-3

Query Match 5.6%; Score 141.2; DB 4; Length 3809;
Best Local Similarity 47.1%; Pred. No. 6.4e-27;
Matches 499; Conservative 1; Mismatches 554; Indels 6; Gaps 2;

QY 1426 ctaacataatgagacaaatccagtgcaacggaagaaacacagtgcttaagct 1485
DB 1933 CTCTTCATCAACGAGAGAAATCCGTGTGAGGCTCTCCAGGAGGCTCCCTTC 1992
QY 1486 gtgttcacagagactgtcttaagagacacagcgag---tggtacgggttcacac 1542
DB 1993 AACTGACGCGAGACTGCTTGCGAGGACGAGAAAGGATCATGAGGAGCCACC 2052
QY 1543 tgcgtcttgaagtggtgcttggtgagccttcccaacaagagtgacctac 1602
DB 2053 TCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2112

QY 1603 agatgccagccttggtggaagaagagtggaacctgaggaagacagacctcccg 1662
DB 2113 GCCTTGAACAGTGGCCAGATGACTTGTGTCACATGAGAACACACCTCTCATTTGCC 2172
QY 1663 cgcacgtggtgttttggttgcttgagacacacctctggtgtgctgagagtaac 1722
DB 2173 AAGGAGATCAGATTTCTGCTGAGGAGACCTTTGGATGACATCACCTCTTTGCC 2232
QY 1723 acgtgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1782
DB 2233 GTCCTGGGAGTTTTCCTGACAGCCTTGTGCTGGGTGTGTTTCAAGTTCCGACACA 2292
QY 1783 cctgtgtgagtgcaagagggcgctgtcttcttcttcttcttcttcttcttctt 1842
DB 2293 CCCATTGTCAAGGCCAACACACGAGAGCTCTTACTCTCTCTCTCTCTCTCTCTCT 2352
QY 1843 gtagtggagcctctatgctcttcttggaaccccaagagcctgctgtgctagc 1902
DB 2353 TCTTTCTCAAGCTCCCTGTTCTTCATCGGAGACCCAGACTGAGCGTCCGCTGCGC 2412
QY 1903 cagacctcttgcccttggtttacacatcttctgctgctgacagtgctcatc 1962
DB 2413 CACCGGCGCTTGGCATACAGCTTCGTGCTGATCATGATCTGTTGTAAGTGAACCAAC 2472
QY 1963 caactaatcatcttcaagtttcccaagagtaactacattaccagcctggtc 2022
DB 2473 CGTGTCTCTCTGCTGTTGAGG---CCAAGATCCCAACAGCTTCCACCGCAAGTGTG 2529
QY 2023 caaagcagcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2082
DB 2530 GGGCTAACCTGACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2589
QY 2083 ctaacttgctggtggtggtggtggtggtggtggtggtggtggtggtggtggt 2142
DB 2590 GTGATCTGCTTACACCGGCGCCCTCAAGCTACCGACAGAGCTGAGATGAG 2649
QY 2143 ctgtgtatgcttgatgacagagagacacactccttggttcttcttcttcttct 2202
DB 2650 ATCATCTTATCATCGTGCACGAGGCTCCCTCATGAGGCTGCTGATGAGCTGAG 2709
QY 2203 aatgctcctctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2262
DB 2710 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2769
QY 2263 aactacaagagcccaatggttcaacttcaacttcaacttcaacttcaacttcaact 2322
DB 2770 AACTCAATGAGAGCCAAAGTTCAATCACTTCAAGATGCTCATGCTTTCATGCTG 2829
QY 2323 gcttcttcaacacagcagcgctctcagagagagtaactcctgctgctgctgctg 2382
DB 2830 TCTTTATTCACAGCCATGACACACACCTATGCAAGTTGTCTCTCCGTAGAGGTAT 2889
QY 2383 gctgggtgagagcctgagcagcgctcctggtggttcttctgctgctgctgctg 2442
DB 2890 GCCATCTTGGACACAGCTTTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2949
QY 2443 atctctgcgcagcctcaacagcagcagcagcagcagcagcagcagcagcagcagc 2482
DB 2950 ATTCTTCAAGCATTCCGCAACACATCAAGAGAGAGTGC 2989

Search completed: June 11, 2002, 06:49:19
Job time: 6410 sec

PF 28-MAR-2001; 2001WO-US09996.
XX
PR 28-MAR-2000; 2000US-192978P.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
PI Zambrowicz B, Sands AT;
XX
DR P-PSDB; AAE11969.
XX
XX MPI: 2001-616474/71.
XX
XX Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
XX
PS Claim 1; Page 71-72; 80pp; English.
XX
XX The present sequence is a cDNA encoding human novel G-protein coupled
CC receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heart beat rate)
CC or behavioral disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.
XX
XX
SQ Sequence 2526 BP; 469 A; 770 C; 697 G; 587 T; 3 other:

DB 421 cttgcagtgatggcctgacagaccacacgctgctgcacacacagccgctctgaagc 480
DB 421 cttgcagtgatggcctgacagaccacacgctgctgcacacacagccgctctgaagc 480
QY 481 cctttccctggtcccatgattgactatgctgcgcagagagagacgtcagctgaagcgg 540
DB 481 cctttccctggtcccatgattgactatgctgcgcagagagagacgtcagctgaagcgg 540
QY 541 cagatccctcttctctgacccatcccatatgacaagtacaaagtggagaccatgctg 600
DB 541 cagatccctcttctctgacccatcccatatgacaagtacaaagtggagaccatgctg 600
QY 601 ctgctgctcagaagttcgggtgagaccgtgacatctctctgtgttgacagcagtaagactat 660
DB 601 ctgctgctcagaagttcgggtgagaccgtgacatctctctgtgttgacagcagtaagactat 660
QY 661 gggcagctaggggtgtaggacactgtagaacccagccagctgctcagggatctgctatgct 720
DB 661 gggcagctaggggtgtaggacactgtagaacccagccagctgctcagggatctgctatgct 720
QY 721 ttcaagacatcatgacctctctctgcccaggtgagcgatgagagatgacgtcctcatg 780
DB 721 ttcaagacatcatgacctctctctgcccaggtgagcgatgagagatgacgtcctcatg 780
QY 781 cgcacactgcccagagccggggccacacgctcgtgtgtttttccagccggagcttgagcc 840
DB 781 cgcacactgcccagagccggggccacacgctcgtgtgtgtttttccagccggagcttgagcc 840
QY 841 aggtgttttttgagtcggtgctgctgacacacgtgacagaggtgaggtgctgctccta 900
DB 841 aggtgttttttgagtcggtgctgctgacacacgtgacagaggtgaggtgctgctccta 900
QY 901 gaagctgggacctctctccagagacatcatctgggtgcccggagatccagctgagatg 960
DB 901 gaagctgggacctctctccagagacatcatctgggtgcccggagatccagctgagatg 960
QY 961 gtgctggtgctggtccatccagaagaggtgtccctgtgacctgaagcgctttaagaagcc 1020
DB 961 gtgctggtgctggtccatccagaagaggtgtccctgtgacctgaagcgctttaagaagcc 1020
QY 1021 tatgcccgggacagacagagagcccttaaggcctttgcccacaagggcctcgtgtcagcagc 1080
DB 1021 tatgcccgggacagacagagagcccttaaggcctttgcccacaagggcctcgtgtcagcagc 1080
QY 1081 aatcagctctgcagagaaatgccaagctttcatgtrcacacagatgcccagaactcaaacg 1140
DB 1081 aatcagctctgcagagaaatgccaagctttcatgtrcacacagatgcccagaactcaaacg 1140
QY 1141 ttctccatgattctgctcctacaacgcatccgggtgtgtatgctgtggtccatggtctc 1200
DB 1141 ttctccatgattctgctcctacaacgcatccgggtgtgtatgctgtggtccatggtctc 1200
QY 1201 caccagctcctgggctgtgctctctgagctgttccagaggcccgagttcaacctggtcgag 1260
DB 1201 caccagctcctgggctgtgctctctgagctgttccagaggcccgagttcaacctggtcgag 1260
QY 1261 ctttggagcagatccacaaggtgtgacttccctctcaacaagagacatgttgagctttaat 1320
DB 1261 ctttggagcagatccacaaggtgtgacttccctctcaacaagagacatgttgagctttaat 1320
QY 1321 gacaaagagatccctctcagtagctataacataattgctggtgagctggtgaatggaccag 1380
DB 1321 gacaaagagatccctctcagtagctataacataattgctggtgagctggtgaatggaccag 1380
QY 1381 ttgacctcagagcttctctggttccctccatggtctcaggttccagctgaactaaatggag 1440
DB 1381 ttgacctcagagcttctctggttccctccatggtctcaggttccagctgaactaaatggag 1440
QY 1441 accaaatccagttggcagcgaagagacacacagctgctgaagtctgtgttccagcgagc 1500
DB 1441 accaaatccagttggcagcgaagagacacacagctgctgaagtctgtgttccagcgagc 1500
QY 1501 tgtcttgaaggcaccagcagatgtgttaacggttccatcaactgtcttgaaggtgtg 1560
DB 1501 tgtcttgaaggcaccagcagatgtgttgaacggttccatcaactgtcttgaaggtgtg 1560

Db 1501 tgtcttgaaaggacacagagtggttacggtttccatcactgctgttgatgtgtg 1560
QY 1561 cccgttgagggtcttggaacttccaaagaagtacacctacagatgcaacgttgagg 1620
Db 1561 cccgttgagggtcttggaacttccaaagaagtacacctacagatgcaacgttgagg 1620
QY 1621 aaagaagagtggtgacactgaaaggaaagcagacgttccgcgcacgttggttttg 1680
Db 1621 aaagaagagtggtgacactgaaaggaaagcagacgttccgcgcacgttggttttg 1680
QY 1681 gcttggtgagacacactcttggtgtgctgtgcaactaaacagctgtgtgtctgtg 1740
Db 1681 gcttggtgagacacactcttggtgtgctgtgcaactaaacagctgtgtgtctgtg 1740
QY 1741 ctgcttggaactgtgtgctgtgttggtgcaactaaacacccctgtgtgaagtgcaaca 1800
Db 1741 ctgcttggaactgtgtgctgtgttggtgcaactaaacacccctgtgtgaagtgcaaca 1800
QY 1801 gggggccgctgtgtcttctatgtgtggtcctccctgagcagagtgatgagcactctat 1860
Db 1801 gggggccgctgtgtcttctatgtgtggtcctccctgagcagagtgatgagcactctat 1860
QY 1861 ggtctcttgagggaacccaagaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1920
Db 1861 ggtctcttgagggaacccaagaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1920
QY 1921 ggttccacatctctgtctgt 1980
Db 1921 ggttccacatctctgtctgt 1980
QY 1981 aagtttccacaagtgtaaccatacttaccacagcctgtgtgtgtgtgtgtgtgtgtgtgt 2040
Db 1981 aagtttccacaagtgtaaccatacttaccacagcctgtgtgtgtgtgtgtgtgtgtgtgt 2040
QY 2041 ctgtttgtatgatcatcagctcagcgccagcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2100
Db 2041 ctgtttgtatgatcatcagctcagcgccagcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2100
QY 2101 tggagcccaactgt 2160
Db 2101 tggagcccaactgt 2160
QY 2161 acaagaagcaactccctgt 2220
Db 2161 acaagaagcaactccctgt 2220
QY 2221 agtgccttgcctgcagctacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2280
Db 2221 agtgccttgcctgcagctacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2280
QY 2281 tgytcaacttgcagctgt 2340
Db 2281 tgytcaacttgcagctgt 2340
QY 2341 agcttctacagcggcaagtactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2400
Db 2341 agcttctacagcggcaagtactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2400
QY 2401 agcagcggtctcggttggt 2460
Db 2401 agcagcggtctcggttggt 2460
QY 2461 ctcaacagacagagcaatttccaggtcctcaatcaagaagcagagcggtcggtgtgtgt 2520
Db 2461 ctcaacagacagagcaatttccaggtcctcaatcaagaagcagagcggtcggtgtgtgt 2520
QY 2521 acctga 2526
Db 2521 acctga 2526

RESULT 2
AAD19505

ID AAD19505 standard; DNA; 2951 BP.
AC AAD19505;
XX 18-DEC-2001 (first entry)
DT Human novel G-protein coupled receptor DNA with 5' and 3' regions.
DE Human: G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW antisense therapy; signal transduction; behavioural disorder; obesity;
KW heartbeats rate; inflammation; immune disorder; diabetes; cancer;
KW coronary disease; ds.
OS Homo sapiens.
XX WO200172842-A2.
XX 04-OCT-2001.
XX 28-MAR-2001; 2001WO-US09996.
XX 28-MAR-2000; 2000US-192978P.
XX (LEXI-) LEXICON GENETICS INC.
XX Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
XX Zambrowicz B, Sands AT;
XX WPI: 2001-616474/71.
XX Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
XX Disclosure: Page 79-80; 80pp; English.
XX The present sequence is human novel G-protein coupled receptor (NGPCR)
CC DNA with 5' and 3' regions. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.
XX Sequence 2951 BP: 559 A; 888 C; 831 G; 669 T; 4 other:
SO

Query Match 100.0%; Score 2524.8; DB 22; Length 2951;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgcctctgacagcgtcgt 60
Db 232 atgtgcctctgacagcgtcgt 291
QY 61 ttgtgcctgacatagcagagtgcttctctacttccacccctcccgagatattactctgt 120
Db 292 ttgtgcctgacatagcagagtgcttctctacttccacccctcccgagatattactctgt 351
QY 121 gaagcctgttccctctccatctgt 180
Db 352 gaagcctgttccctctccatctgt 411

[illegible]

QY	1261	cttttgagcagatccatcaaaagtgatattcttcttcaacaaagagacacgttgcgtttaat	1320
Db	1492	cttttttgagcagatccatcaaaagtgatattcttcttcaacaaagacacgttgcgtttaat	1551
QY	1321	gacaaacagatccccctcagttagatataaactatgttccttgagacttgaaatggacccaag	1380
Db	1552	gacaaacagatccccctcagttagatataaactatgttccttgagacttgaaatggacccaag	1611
QY	1381	tggaccttaocggtccctcgtgtctctccacaatggtctccagttcagctaaacataatgag	1440
Db	1612	tggaccttaocggtccctcgtgtctctccacaatggtctccagttcagctaaacataatgag	1671
QY	1441	accaaaatccagttgacagagaaagacaacaagtgctcctaagtctgtgtgttcaagcgac	1500
Db	1672	accaaaatccagttgacagagaaagacaacaagtgctcctaagtctgtgtgttcaagcgac	1731
QY	1501	tgtcttgaaagggcacacagcagagtgtgttaacgggtttccatacctgtctgtttgaaftgttg	1560
Db	1732	tgtcttgaaagggcacacagcagagtgtgttaacgggtttccatacctgtctgtttgaaftgttg	1791
QY	1561	ccctgttgagggtctgagaccttccctcaacaagaagtgaacctctacaagatgcacaagcctgtggg	1620
Db	1792	ccctgttgagggtctgagaccttccctcaacaagaagtgaacctctacaagatgcacaagcctgtggg	1851
QY	1621	aaagaagaagtgtgacacccctgagaggaaagccagacacctgtccctccgagacgttgtgttttg	1680
Db	1852	aaagaagaagtgtgacacccctgagaggaaagccagacacctgtccctccgagacgttgtgttttg	1911
QY	1681	gctttgcttgaacaacacctctgtgggtgtctgtctgagagctaaacaacgtctctgtctgtctg	1740
Db	1912	gctttgcttgaacaacacctctgtgggtgtctgtctgagagctaaacaacgtctctgtctgtctg	1971
QY	1741	ctgcttggagactgtgacctgttctgtctgacactaagacacccctgtgggtgaagtcaagca	1800
Db	1972	ctgcttggagactgtgacctgttctgtctgacactaagacacccctgtgggtgaagtcaagca	2031
QY	1801	ggggggcgcgcgtgtgttttcttatagttcttgagctccctccggagaagaagtgagggagacctcat	1860
Db	2032	ggggggcgcgcgtgtgttttcttatagttcttgagctccctccggagaagaagtgagggagacctcat	2091
QY	1861	ggcttcttcttgaggaaacccaagaagsgctctgtgtctgtgcacagccagccctctttggccct	1920
Db	2092	ggcttcttcttgaggaaacccaagaagsgctctgtgtctgtgcacagccctctttggccct	2151
QY	1921	ggttcaaccatctccctgtctcgtctgacagtttgctcatctcaatccaatcatcatcttc	1980
Db	2152	ggttcaaccatctctcgtctcgtctgacagtttgctcatctcaatccaatcatcatcttc	2211
QY	1981	aagtttccaacaaagatgaactcatcatcttcaacacgctctgggtgtccaaaacacaggtgtctgc	2040
Db	2212	aagtttccaacaaagatgaactcatcatcttcaacacgctctgggtgtccaaaacacaggtgtctgc	2271
QY	2041	ctgtttgtgatbatatagctcagcggccccaagctgctatactgttctcaacttggctgtgtgtg	2100
Db	2272	ctgtttgtgatbatatagctcagcggccccaagctgctatactgttctcaacttggctgtgtgtg	2331
QY	2101	tggacccaactgtcctgtcagtgaaagatatcacagcgcttccccaactgtgtgtatgtatgtg	2160
Db	2332	tggacccaactgtcctgtcagtgaaagatatcacagcgcttccccaactgtgtgtatgtatgtg	2391
QY	2161	aaagaagacaactccctgtgggtctatactagtcttccctctacaatgagcctctctcatc	2220
Db	2392	aaagaagacaactccctgtgggtctatactagtcttccctctacaatgagcctctctcatc	2451
QY	2221	agtgccttgcgcctgagactaacctgtgttaaggaactgtgcagaagaactaaagaaggcca	2280
Db	2452	agtgccttgcgcctgagactaacctgtgttaaggaactgtgcagaagaactaaagaaggcca	2511
QY	2281	tgtgtcaccttaagcgtctgtctcaaatctgtgtctcctgtgatcgccttcttcaacaagcgc	2340
Db	2512	tgtgtcaccttaagcgtctgtctcaaatctgtgtctcctgtgatcgccttcttcaacaagcgc	2571
QY	2341	agcgtctacagagcgaaatlaactcgtctgcgcggccaacatgatgtgctggctgtgaagacgctg	2400

```

|||||
Db 2572 agcgctacagcagcgaatgactgcctgcggcacaacatgctgctgagcagcctg 2631
Oy 2401 agcagcgctctcggtgggtattttctgctaagtctacgfatcctctgcgcccacac 2460
Db 2632 agcagcgctctcggtgggtattttctgctaagtctacgfatcctctgcgcccacac 2691
Oy 2461 ctcaacacacagacgactctccagcctccattcagagctaacagagcgctgcgctcc 2520
Db 2692 ctcaacacacagacgactctccagcctccattcagagctaacagagcgctgcgctcc 2751

Oy 2521 acctga 2526
Db 2752 acctga 2757

RESULT 3
AADI7516
ID AADI7516 standard; cDNA; 2526 BP.
XX
AC AADI7516;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human taste receptor, hTIR1 cDNA coding sequence.
XX
KW Human; taste-cell-specific G protein-coupled receptor; hTIR1; drug;
KW genetic modulation; pharmaceutical; taste sensation; food industry;
KW chemosensory transduction; ss.
XX
OS Homo sapiens.
XX
FH Key 1..2526 location/Qualifiers
FT CDS /*tag= a
FT /product= "Human taste receptor, hTIR1 protein"
FT /trans_except= (Pos:820..823, aa:Phe)
FT /note= "This codon has an apparent 1 nucleotide insertion
FT which alters the reading frame"
XX
XX WO200166563-A2.
XX
XX 13-SEP-2001.
XX
XX 07-MAR-2001; 2001WO-US07265.
XX
XX 07-MAR-2000; 2000US-0187546.
XX
XX 07-APR-2000; 2000US-0195536.
XX
XX 06-JUN-2000; 2000US-0209840.
XX
XX 23-JUN-2000; 2000US-0214213.
XX
XX 17-AUG-2000; 2000US-0226448.
XX
XX 03-JAN-2001; 2001US-0259227.
XX
XX (SENO-) SENOMYX INC.
XX
XX Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
XX WPI: 2001-583267/65.
XX DR P-PSDB; AAEI0372.
XX
XX New mammalian taste-cell-specific G protein-coupled receptor
XX polypeptides for identifying compounds that modulate taste signaling
XX are useful in food, to modulate the sweet taste of foods or drugs -
XX
XX Claim 1; Page 82-83; 119pp: English.
XX
XX The invention relates to mammalian taste-cell-specific G protein-coupled
XX receptors, TIR and their corresponding cDNA molecules. Taste receptors,
XX TIR are useful for screening compounds which are used to activate or
XX modulate chemosensory transduction, such as taste sensation. The
XX identification and isolation of novel taste receptors and taste
XX signalling molecules allow for new methods of chemical and genetic
XX modulation of taste transduction pathways. The taste modulating

```

CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. TIRs are also useful in
 CC biochemical assay for identifying tastant (TIR) ligands having binding
 CC specificity for TIR involved in taste signalling. The present cDNA
 CC sequence is human taste-cell-specific G protein-coupled receptor, hTIR1
 CC coding sequence.

Sequence 2526 BP; 470 A; 770 C; 697 G; 589 T; 0 other;

Query Match 99.0%; Score 2500.2; DB 22; Length 2526;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2521; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

```

Oy 1 atgtctcttcagcagcctgcctgctggtgcctgcacagcttcttctctgctgcgc 60
Db 1 atgtctcttcagcagcctgcctgctggtgcctgcacagcttcttctctgctgcgc 60
Oy 61 ttgctctccatagacagcagatctctctctgacttccacctcccgagattactctctg 120
Db 61 ttgctctccatagacagcagatctctctctgacttccacctcccgagattactctctg 120
Oy 121 gcaagcctgttcctctccatctgtctgtctgcaggtgagacagaccgaggtgac 180
Db 121 gcaagcctgttcctctccatctgtctgtctgcaggtgagacagaccgaggtgac 180
Oy 181 ctgtgtgacaggtctgttagcttcaatagacatggtctacacaccttccaggtatcg 240
Db 181 ctgtgtgacaggtctgttagcttcaatagacatggtctacacaccttccaggtatcg 240
Oy 241 ctgtgtgtgagagataaacaacacccgcgcgtgcacacataaccttggtgac 300
Db 241 ctgtgtgtgagagataaacaacacccgcgcgtgcacacataaccttggtgac 300
Oy 301 cagctgtatgattgtgttctgtactctgcacatgtgtatgctgcacgctgagagtgctcc 360
Db 301 cagctgtatgattgtgttctgtactctgcacatgtgtatgctgcacgctgagagtgctcc 360
Oy 361 ctgcagagggcaaacacacatagatctccaaggaacaccttccactatccctcagtg 420
Db 361 ctgcagagggcaaacacacatagatctccaaggaacaccttccactatccctcagtg 420
Oy 421 ctggaagtgtatggcctgcagacacacacacgctgtctgcacacacacgcccctgtgagc 480
Db 421 ctggaagtgtatggcctgcagacacacacacgctgtctgcacacacacgcccctgtgagc 480
Oy 481 ccttctctgtgcccattagctatgctgcgacagcagcagctcagcgttgagacgg 540
Db 481 ccttctctgtgcccattagctatgctgcgacagcagcagcagctcagcgttgagacgg 540
Oy 541 cagtatccctcttctctctgcacacatccccaatgacagatgagagacatgtg 600
Db 541 cagtatccctcttctctctgcacacatccccaatgacagatgagagacatgtg 600
Oy 601 ctgctgcgcagaaagtctgggtgacactgacatctctgtgtggagcagatgacacat 660
Db 601 ctgctgcgcagaaagtctgggtgacactgacatctctgtgtggagcagatgacacat 660
Oy 661 gggcagctaggggtgcaggaacttgagaacacagcgcactggtcagggatctgcatgt 720
Db 661 gggcagctaggggtgcaggaacttgagaacacagcgcactggtcagggatctgcatgt 720
Oy 721 ttcaagagacatcatgctcttctctctccaggtggcagatgagagatgcaggtcctatg 780
Db 721 ttcaagagacatcatgctcttctctctccaggtggcagatgagagatgcaggtcctatg 780
Oy 781 cggccacctggccacagcgggcgccacgctgtggtg-cttttccacccgacattggc 839
Db 781 cggccacctggccacagcgggcgccacgctgtggtg-cttttccacccgacattggc 840
Oy 840 cagggtgttttgcagctcgtgtgtctgacacacactgacatgcaaggtgtgtgctgcctc 899
Db 840 cagggtgttttgcagctcgtgtgtgtctgacacacactgacatgcaaggtgtgtgctgcctc 899

```

Db 841 caaggtgttttttcagagtcggtgctgacaaactgactgcaaggtgtggtgcctc 900
QY 900 aagaagccttgagcccttcacagcacatcaactgaggtgccccgagatccagcgcatgtggat 959
Db 901 aagaagccttgagcccttcacagcacatcaactgaggtgccccgagatccagcgcatgtggat 960
QY 960 ggtgtcgtggtggtgacatccagaagaaggtgtccctggtcctgaagggcgtttgaagaagc 1019
Db 961 ggtgtcgtggtggtgacatccagaagaaggtgtc -cctggtcctgaagggcgtttgaagaagc 1019
QY 1020 ctatgcccgggacagaagaagggccctcctgagcccttcgaagaagggctccgggtgagcgag 1079
Db 1020 ctatgcccgggacagaagaagaagggccctcctgagcccttcgaagaagggctccgggtgagcgag 1079
QY 1080 caatcagcctctgacagaagaatgccaagcttcatcagcacacagatgcccagaagtcacaagc 1139
Db 1080 caatcagcctctgacagaagaatgccaagcttcatcagcacacagatgcccagaagtcacaagc 1139
QY 1140 cttctccatgaggttctgctctac 1199
Db 1140 cttctccatgaggttctgctctac 1199
QY 1200 ccaacagcctctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1259
Db 1200 ccaacagcctctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1259
QY 1260 gcttttgagagacatccacaagaatgcatctccttctacacaagaagacacatggtggtttaa 1319
Db 1260 gcttttgagagacatccacaagaatgcatctccttctacacaagaagacacatggtggtttaa 1319
QY 1320 tgacaacaagaatccctcctcagtagcataaataatgctcgtgagatggaatggaaccaa 1379
Db 1320 tgacaacaagaatccctcctcagtagcataaataatgctcgtgagatggaatggaaccaa 1379
QY 1380 gtggaacttcaacgtctcgtctcctcacaatgctcaggtcagctaaacataatga 1439
Db 1380 gtggaacttcaacgtctcgtctcctcacaatgctcaggtcagctaaacataatga 1439
QY 1440 gaccaaataccagtgagcagcgaagaagaacaacaggtgcttaagtcgtgtgtccagcga 1499
Db 1440 gaccaaataccagtgagcagcgaagaagaacaacaggtgcttaagtcgtgtgtccagcga 1499
QY 1500 cgtctctgaagggacacagcaggtgttaacgggttccatcactgctgttgaagtgtgt 1559
Db 1500 cgtctctgaagggacacagcaggtgttaacgggttccatcactgctgttgaagtgtgt 1559
QY 1560 ggcctgtgagggagcagccttcccaacaagaatgagcctcctaagaatgccaagcctgtgtg 1619
Db 1560 ggcctgtgagggagcagccttcccaacaagaatgagcctcctaagaatgccaagcctgtgtg 1619
QY 1620 gaaagaagaatgagcagcctgaggaagaagcagacatgctcccgacgtgtgtgtttt 1679
Db 1620 gaaagaagaatgagcagcctgaggaagaagcagacatgctcccgacgtgtgtgtttt 1679
QY 1680 ggcctgtgagggacacacacttgggtgtgctgtgagcagcctaagcgtctgtctgtct 1739
Db 1680 ggcctgtgagggacacacacttgggtgtgctgtgagcagcctaagcgtctgtctgtct 1739
QY 1740 gctgtctgagggagcgtggt 1799
Db 1740 gctgtctgagggagcgtggt 1799
QY 1800 aagggagcgcctgt 1859
Db 1800 aagggagcgcctgt 1859
QY 1860 tggctctcttggggaaacccaagaagcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1919
Db 1860 tggctctcttggggaaacccaagaagcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1919
QY 1920 tgggttccacatcttctcctgctgagcaggtgctatccatccataataataatctt 1979
Db 1920 tgggttccacatcttctcctgctgagcaggtgctatccatccataataataatctt 1979

QY 1980 caagttttccaccaaggttacctacatctac 2039
Db 1980 caagttttccaccaaggttacctacatctac 2039
QY 2040 cctgttttgatgatcatcagctcagcggccagcgtctatctgtctactgtgtgtgtgtgtgtgtgtgtgtgt 2099
Db 2040 cctgttttgatgatcatcagctcagcggccagcgtctatctgtctactgtgtgtgtgtgtgtgtgtgtgtgt 2099
QY 2100 gtgagcccaactcgtctgtaggaatatacagcgtctcccccactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2159
Db 2100 gtgagcccaactcgtctgtaggaatatacagcgtctcccccactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2159
QY 2160 cacagagccaactccctggtgtcttactactgt 2219
Db 2160 cacagagccaactccctggtgtcttactactgt 2219
QY 2220 cagtgcttctgctgtgagcctggtgtaggagcgt 2279
Db 2220 cagtgcttctgctgtgagcctggtgtaggagcgt 2279
QY 2280 atgtgtcacctcagcctgtcttcaacttcgt 2339
Db 2280 atgtgtcacctcagcctgtcttcaacttcgt 2339
QY 2340 cagcgtctacagcagcaggtacacgtcgt 2399
Db 2340 cagcgtctacagcagcaggtacacgtcgt 2399
QY 2400 gtagcagcgtcttctggt 2459
Db 2400 gtagcagcgtcttctggt 2459
QY 2460 cctcaacagc 2519
Db 2460 cctcaacagc 2519
QY 2520 cacttga 2526
Db 2520 cacttga 2526

RESULT 4
AADI9502
ID AADI9502 standard; cDNA; 2292 BP.
XX
AC AADI9502;
DT 18-DEC-2001 (first entry)
DE Human novel G-protein coupled receptor (NGPCR) cDNA #2.
XX
KW Human: G-protein coupled receptor; GPCR; gene therapy; drug screening;
antisense therapy; signal transduction; behavioural disorder; obesity;
heartbeats rate; inflammation; immune disorder; diabetes; cancer;
coronary disease; ss.
XX
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FT CDS 1..2292
FT /tag= a
FT /product= "Human NGPCR protein"
FT /tag= b
FT /note= "Encodes Ser"
FT /tag= c
FT /note= "Encodes Ala"
XX
PN MO200172842-A2.
XX
PD 04-OCT-2001.

QY	550	tcltctctgagcacatccccaatgcagaagtacaaagtctgagaaacatgctgtctgctg	609
Db	358	tcltctctctgagcacatccccaatgcagaagtacaaagtctgagaaacatgctgtctgctg	417
QY	610	cagaagaatctcggtctgagacctgtgatactctctctggttttgacagaagtgagacatatactgagacata	669
Db	418	cagaagaatctcggtctgagacctgtgatactctctctggttttgacagaagtgagacatatactgagacata	477
QY	670	ggggctgacagagcacatgtagaagcaacacagagccactgtgttcaagagatctatgcatgtcttcaagagac	729
Db	478	ggggctgacagagcacatgtagaagcaacacagagccactgtgttcaagagatctatgcatgtcttcaagagac	537
QY	730	atcaatgacctctctctcccaagctgtggctgatactgagagagatgacgtgccatactgcgcaactg	789
Db	538	atcaatgacctctctctcccaagctgtggctgatactgagagagatgacgtgccatactgcgcaactg	597
QY	790	ggccaaagcccgggggccaaacgctcgtgtgtgttttttccagaacggcaattggccaaagtggttt	849
Db	598	ggccaaagcccgggggccaaacgctcgtgtgtgttttttccagaacggcaattggccaaagtggttt	657
QY	850	ttcagatctcgtctgtctgcagaaacaaacccctgagctgtgcaagagtctgtggtctgcgcttcaagaaagctg	909
Db	658	ttcagatctcgtctgtctgcagaaacaaacccctgagctgtgcaagagtctgtggtctgcgcttcaagaaagctg	717
QY	910	ggcctctcccaagtgacatactgctgggtgtgcccgagatccaaagcattgtggatgtgtctgtgc	969
Db	718	ggcctctcccaagtgacatactgctgggtgtgcccgagatccaaagcattgtggatgtgtctgtgc	777
QY	970	gttgagccatcccaagaagagagctgtccctctgcgtctgaaagagcgttttgaaagaagccatactgcccgg	1029
Db	778	gttgagccatcccaagaagagagctgtccctctgcgtctgaaagagcgttttgaaagaagccatactgcccgg	837
QY	1030	gcacaaacaaagagagggccctctaaagcctctgcacaaagagggctccctgtgtgacagacgaatacagctc	1089
Db	838	gcacaaacaaagagagggccctctaaagcctctgcacaaagagggctccctgtgtgacagacgaatacagctc	896
QY	1090	tgccagagaatgccaagaatcttcatactgacacacagaatgcccagaagctcaagaagctctctccatg	1149
Db	897	tgccagagaatgccaagaatcttcatactgacacacagaatgcccagaagctcaagaagctctctccatg	956
QY	1150	agttctctcctaaacagacatacgcgggctgtgtatgctgtggccatctgagcctccaaacagctc	1209
Db	957	agttctctcctaaacagacatacgcgggctgtgtatgctgtggccatctgagcctccaaacagctc	1016
QY	1210	ctggggctgtgctctctctgtagagctctgttccaaaggggcagagcttcaaccccttgagcagcttttgag	1269
Db	1017	ctggggctgtgctctctctgtagagctctgttccaaaggggcagagcttcaaccccttgagcagcttttgag	1076
QY	1270	cagatcccaacaaagtgcattctctctcttactacaaagagacactgtgtgcgttttaatgtgacaacaga	1329
Db	1077	cagatcccaacaaagtgcattctctctcttactacaaagagacactgtgtgcgttttaatgtgacaacaga	1136
QY	1330	gataccctcctaaagtataataataattgctcctgtagactgtagaattgtagagcccaagtgtgagaccttc	1389
Db	1137	gataccctcctaaagtataataataattgctcctgtagactgtagaattgtagagcccaagtgtgagaccttc	1196
QY	1390	acggctctcctggtttctccatactgctcccaagtctccagctcagctaaataataatgtgagaccaaatc	1449
Db	1197	acggctctcctggtttctccatactgctcccaagtctccagctcagctaaataataatgtgagaccaaatc	1256
QY	1450	cagtgtgcacagaaagagacaacacagagctgtcctaagatctgtgtgttccagagcagctgttctgaa	1509
Db	1257	cagtgtgcacagaaagagacaacacagagctgtcctaagatctgtgtgttccagagcagctgttctgaa	1316
QY	1510	gggacacagcagagtgtgtttaaagggtttccatacactgtcgtcttgagatgtgtgtgcctgtggg	1569
Db	1317	gggacacagcagagtgtgtttaaagggtttccatacactgtcgtcttgagatgtgtgtgtgcctgtggg	1376
QY	1570	gcttgagaccttccctcaacacagagtgagcctctactaagaatgcacagcctgtgtgtggaaagaaag	1629
Db	1377	gcttgagaccttccctcaacacagagtgagcctctactaagaatgcacagcctgtgtgtggaaagaaag	1436
QY	1630	tgaggacaccttgagaggaagccagaagcctgtcttcccgcgacactgtgtgttttctgtcttggt	1689

Db	1437	ttgggaacctgaggggaagccagacctgccttcgcggcaacgtgagtgatttcttgcttgctgct	1496
Qy	1690	gagcaacacctctcttggtggtcgtcgtcgagcgtcaaacagcctctctgctcgtcgtcttggtg	1749
Db	1437	gagcaacacctctcttggtggtcgtcgtcgagcgtcaaacagcctctctgctcgtcgtcgtcttggtg	1556
Qy	1750	actgctgcgtcgttcttgctcgtcgcaactagacaacccctctgtgtgtagagtcagcagggggccgc	1809
Db	1557	actgctgcgtcgttcttgctcgtcgcaactagacaacccctctgtgtgtagagtcagcagggggccgc	1616
Qy	1810	ctgtgcttcttctatgcgtcgggctccctcgtgcagcaggtgagtgagcagcctctatagctctctt	1869
Db	1617	ctgtgcttcttctatgcgtcgggctccctcgtgcagcaggtgagtgagcagcctctatagctctctt	1676
Qy	1870	gggggaaccacaagagccttgctgtctgtctacgcgaagggcctcttttgcccttgatttcaac	1929
Db	1677	gggggaaccacaagagccttgctgtctgtctacgcgaagggcctcttttgcccttgatttcaac	1736
Qy	1930	atctctcgtctcgtcgtcgtacagtttgctgtcatcttccaactaatcatcatcttcaagtcttcc	1989
Db	1737	atctctcgtctcgtcgtcgtacagtttgctgtcatcttccaactaatcatcatcttcaagtcttcc	1766
Qy	1990	accaaggtacctaatattctacacagcctctgtgttccaacacaggtgctggtcgtgtttgtg	2049
Db	1797	accaaggtacctaatattctacacagcctctgtgttccaacacaggtgctggtcgtgtttgtg	1856
Qy	2050	atgatacagctacagcggcccgacgtcgtctatctgtctctaaccttggtcgtgtgtgtagaccca	2109
Db	1857	atgatacagctacagcggcccgacgtcgtctatctgtctctaaccttggtcgtgtgtgtagaccca	1916
Qy	2110	ctgcctcgtctagggataaccagcgcttcccccactctgtgtgtgtagcttgagtgacagagacc	2169
Db	1917	ctgcctcgtctagggataaccagcgcttcccccactctgtgtgtgtagcttgagtgacagagacc	1976
Qy	2170	aactccctgggctcttcaactgagcggcttccctctataatgagcgtcctctctcatcagtgacctt	2229
Db	1977	aactccctgggctcttcaactgagcggcttccctctataatgagcgtcctctctcatcagtgacctt	2036
Qy	2230	gctctgcaactcactctgtgttgaaagacttgccagagaaactacaacagagccaaatgtygtacc	2289
Db	2037	gctctgcaactcactcctgggttgaaagacttgccagagaaactacaacagagccaaatgtygtacc	2096
Qy	2290	ttaagctcgtcctctcaaatcttcgtgtgctcttgatcgtcgtctctctcaacagcgccagcgtctac	2349
Db	2097	ttaagctcgtcctctcaaatcttcgtgtgctcttgatcgtcgtctctctcaacagcgccagcgtctac	2156
Qy	2350	gaacggcaagtactcgtcctcgtcgccacaatgatatgctgtggcgtgagcagccttgagcagcgcc	2409
Db	2157	gaacggcaagtactcgtcctcgtcgccacaatgatatgactgtggcgtgagcagccttgagcagcgcc	2216
Qy	2410	ttcgtgtggtattttctgcgtctaagtgtagatgtagtctctctgcgcgcacagactcaaacgc	2469
Db	2217	ttcgtgtggtattttctgcgtctaagtgtagatgtagtctctctgcgcgcacagactcaaacgc	2276
Qy	2470	acagagcaactccagcgtcctcatctcagagactaacagagcgctctgagctccacctgtga	2526
Db	2277	acagagcaactccagcgtcctcatctcagagactaacagagcgctctgagctccacctgtga	2333
RESULT	6		
AAZ50744			
ID	AAZ50744	standard; cDNA; 2579 BP.	
XX	AAZ50744;		
AC			
XX			
DT	31-MAY-2000	(first entry)	
XX			
DE		Mouse sensory transduction G-protein coupled receptor-B3 cDNA.	
KM		Mouse; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;	
KW		Sensory cell; taste receptor cell; screen; taste modulator;	
		pharmaceutical; food industry; taste topographic map; tongue; ss.	

XX	Mus sp.	Location/Qualifiers
XX	Key	11..2539
XX	FT CDS	/tag= a
XX	FT	/product= "GPCR-B3"
XX	PN	WO200006592-A1.
XX	PD	10-FEB-2000.
XX	PF	27-JUL-1999; 99WO-US17099.
XX	PR	28-JUL-1998; 98US-0094465.
XX	PA	(REGC) UNITV CALIFORNIA.
XX	XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	P1	Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;
XX	XX	WPI: 2000-205451/18.
XX	DR	P-PSDB; AAV45022.
XX	PT	New isolated sensory transduction G-protein coupled receptor, useful
XX	PT	for developing products for use in studying and modulating the taste
XX	XX	transduction pathway -
XX	PS	Claim 5; Pages 77-79; 83pp; English.
XX	CC	The present sequence is a CDNA encoding a taste cell specific G-protein
XX	CC	coupled receptor, GPCR-B3 which is involved in sensory transduction.
XX	CC	This sequence was obtained from mouse circumvallate and foliate papillae.
XX	CC	GPCR-B3 is specifically expressed in foliate and fungiform cells, with
XX	CC	lower expression in circumvallate taste receptor cells of the tongue.
XX	CC	The present sequence is used to screen compounds that modulate sensory
XX	CC	signalling in taste cells, especially taste modulators useful in
XX	CC	pharmaceutical and food industries to customise taste. The sequence
XX	CC	can also be used as probe for identifying taste cells and
XX	CC	subsets of taste receptor cells such as foliate, fungiform and
XX	CC	circumvallate. Such probes are also useful to generate taste
XX	CC	topographic maps that elucidate the relationship between the taste
XX	CC	cells of the tongue and sensory neurons leading to taste centres
XX	XX	in the brain.
XX	Sequence	2579 BP, 492 A; 771 C; 698 G; 618 T; 0 other;

Query Match	66.9%;	Score 1689.4;	DB 21;	Length 2579;
Best Local Similarity	79.6%;	Pred. No. 0;		
Matches 1993;	Conservative 3;	Mismatches 509;	Indels 0;	Gaps 0

QY	22	ctggtcgcgcctcgaacttcattcttcctcgtctcgtctcgtcgttcgtcctgcatacgcagag	81
DB	35	ctgctcagcctcgaactcgtcgtctctcctcgtcgttcgtcgttcgaatgcgaagagacagaa	94
QY	82	tcttctcgaattacacctcccgagaaattacctcctcgcagagcctgttcctccat	141
DB	95	tctctcgaagtllacagctcctcctcgttggaacttcctcctcgcagagcctgttccctccat	154
QY	142	tctgctctctcgcaggtgagacagaccgcgaagtgcacctgtgtgacaggtcttgcagc	201
DB	155	gctgactgctcgtcaggtgagacacagacctctgtgtgacaagtgtgacaggtctgcagc	214
QY	202	ttcaatgagcattgctacacctcttcacagctatagcgtctggggttaagagataaac	261
DB	215	ttcaagcgcattgctatcactcttcacagcattgcgttllacacglttagagataaac	274
QY	262	aactcaagcgcctctgcgccaacatcaacctggggtgacagctgattatgtgtt	321
DB	275	aactcacagcctcgtctcccaaatcacacctggggtatgaactgtatgaactgtgtgtca	334
QY	322	gacttcgcaatgtgtatgacagcgttgagatgtctcctcctcgcaggggacaacacata	381

[illegible]

[illegible]

QY	1205	agctccctgggctctgagctctcgtgagagcttgcgcagggccgagtcctacccctggcagcttt	1264
Db	1258	agctctcctggagatgtatctctctgtgagatctctgtctccacagagcccgacagctctacccctggcagctttc	1317
QY	1265	tggagcagatccacaaaggtgcatcttctctctacaaagggacatctgtgtgtttaatgaca	1324
Db	1318	ttcagcagatctacaaagtgaatttctctctacatgagaaatactgtgcatttgatgaca	1377
QY	1325	acagagatccctcagtaataataacaaatgtccctgtggagctgtgaaatggacccaagtgtga	1384
Db	1378	acggggagacacctcagatgtactacagacatactgcgcctgtggagctgtgaaatggacccaagtgtga	1437
QY	1385	cccttcacgctctccggtttctctccacagatctcagcttcaagcttcaagcaataaattggacaga	1444
Db	1438	cccttcagatataatctgtgcctctgtccctcactgtctccagttcactctgtgcacataaataagaca	1497
QY	1445	aaatccagttggacacggaaagagacacccaggctggcctaagtctgtgtgtttccagcgagctgc	1504
Db	1498	aaatccagttggacacggagagacacaaatcaggtgtccctgtcagatgtgttaccacgagctgc	1557
QY	1505	ttgaaaggccacacagccaggtgtgttaacgggtttccataactgtctgtgttgatgtgtggccct	1566
Db	1558	tggcagggccacacaaagggtgtgtgtgtgtgtgtttccacacacgtctgtcttggagttgtggccct	1617
QY	1565	gtggggcgtgggagacctctccacaaagaggtgtacactcaagatgtccagagctctgtggaaag	1624
Db	1618	ggagaaagctgggagacctctccacaaatggagtgagtgatgtccacatctgcagagcctgtggaaag	1677
QY	1625	aagagttgggacacctgtgagggaggaagccagacactgtctcccgagcactgtgtgtgttttgctt	1684
Db	1678	aagaaattgggacacccaaaggagagacatactgtctctccacacacaggtgtggagatcttctgtgctt	1737
QY	1685	tggcgtgagacacacactctgtgggtgtgtgtgtgtgacgtctaaacagcgtctgtgcgtgcgtgc	1744
Db	1738	ggcagaaacacatactctcttgggt	1797
QY	1745	ttggagactgtgcgtcgtgtttgtcctgtgcacataagacacccctgtgtgtgtgtgtgtgtgtgtgt	1804
Db	1798	ttggagacgtgcgtgcgtgtttctcctgtgacatttccacacacactgtgagatgtacagctgtggg	1857
QY	1805	ggcgcgcctgtgtcttctctgtatgtgtgtgtgcctctggcacaagatgtgtggcagccctatgtct	1864
Db	1858	gtaagcgtgtgtcttccatcagtcagtcgtgtccctgtgtgcggagagttgtcagctctctatagct	1917
QY	1865	tcttctgggaaacccacaaagcgtcgtgtctgtctacgcacagagccctcttgcctgtgt	1924
Db	1918	tcttccgggagccacaggtgtgcgcgctgtgtctgtgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1977
QY	1925	tcacacatttccctgtctctgtcgtcagtttgcgtcatttccaaactaaatcacatctccaagt	1984
Db	1978	tgtgcacattctctcctcctgcgtcgtcagaaatccgcctctccacacacgtgtcacaatctccaagt	2037
QY	1985	tttcacacaaaggtacatactacatcttcaacagacgcctgtgtgttcaaaaacacaggtgtgtgcgt	2044
Db	2038	tttccacaaaggtgtccacatctacacgtacacccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2097
QY	2045	ttgtgtatgtacagctcagcgccagctgtctatctgtcttaactgtgcgtgtgtgtgtgtga	2104
Db	2098	tcgtcatctgtcagctccacaggttgcatttgcattgtctatctgtctcacaatgtcgtgtgaatgtga	2157
QY	2105	cccacactcctcgtcagggaaataacagagcttccccatctgtgtgtgtgtgtgtgtgtgtgtgtgt	2164
Db	2158	cccacacgcccacaaagggaaataacagagcttccccatctcgtgtgtatcttcgaggtcagag	2217
QY	2165	agacaaatccctcgtgtcttatactgtgccttctctcctacaaatgtgcgcctccctccatcaagt	2224
Db	2218	aggtacactcgtgtaggtctcctgt	2277
QY	2225	cccttgcctgcagctacactcgtgtgtgaagactgtccagagatacacaagagccaaatgtgt	2284
Db	2278	cccttgcctgcagctacactcgtgtgtgaagactgtccagagatacataatgaagccaaatgtgt	2337

Oy 2285 tcacctcagcctgtcttcaacttcgtctgctgacgcctctctcaccagcagcagc 2344
|||||
Db 2338 tcaactcagcctgtctcctcaacttcgtatctctgtacgctcttcaacatggccagca 2397
Oy 2345 tctacagcagcaagtactcgtctgcgcgcacaataatgctggtgctgagcagcctgaagca 2404
|||||
Db 2398 ttaccagagcagcactcgtcgtcgtcgaatgctgtcgcaggtgacacacacagagcg 2457
Oy 2405 ggcgttggtggtggtattcttcgtcctaagtgtactgatactctgcgcgcacagcctca 2464
|||||
Db 2458 ggcgttcagcgttactctccccaagtgctatgtatctctgcgcgtccagaactca 2517
Oy 2465 acagacagagcactcagcagcctcattcagactacagagcgctgcgcctcacct 2524
|||||
Db 2518 acaatacagaaacttcaagcctccatccaggaactacaagagcgctgcgcgaactact 2577
Oy 2525 ga 2526
Db 2578 ga 2579

RESULT 8

AAD19503 standard; cDNA: 1101 BP.

AAD19503;

18-DEC-2001 (first entry)

Human novel G-protein coupled receptor (NGPCR) cDNA #3.

Human; G-protein coupled receptor; GPCR; gene therapy; drug screening; antisense-therapy; signal transduction; behavioural disorder; obesity; heartbeats rate; inflammation; immune disorder; diabetes; cancer;

coronary disease; ss.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

Homo sapiens.

CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.

SQ Sequence 1101 BP; 219 A; 308 C; 323 G; 250 T; 1 other:

Query Match 39.6%; Score 1001.2; DB 22; Length 1101;
Best Local Similarity 99.6%; Pred. No. 3e-243;
Matches 1004; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 595 atgtgtcgtcgtcgtcagaagttcgggtggtgacctggtctctgtgtgagcaggtgac 654

Db 1 atgtgtcgtcgtcgtcagaagttcgggtggtgacctggtctctgtgtgagcaggtgac 60

Oy 655 gactatggcagcagtaggggtgtagcagcagcagcagcagcagcagcagcagcagcagcagc 714

Db 61 gactatggcagcagtaggggtgtagcagcagcagcagcagcagcagcagcagcagcagcagc 120

Oy 715 attgtctcaagaacataatgacctctctcgtcagcagcagcagcagcagcagcagcagcagc 774

Db 121 attgtctcaagaacataatgacctctctcgtcagcagcagcagcagcagcagcagcagcagc 180

Oy 775 ctcatgcccacactgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 834

Db 181 ctcatgcccacactgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240

Oy 835 ttggccagcagcgtttttcgaagcgtgtgtgacctgacctgacctgacctgacctgacctgacct 894

Db 241 ttggccagcagcgtttttcgaagcgtgtgtgacctgacctgacctgacctgacctgacctgacct 300

Oy 895 gctcagaagccttgagcctctcagcagcagcagcagcagcagcagcagcagcagcagcagc 954

Db 301 gctcagaagccttgagcctctcagcagcagcagcagcagcagcagcagcagcagcagcagc 360

Oy 955 gggatgt 1014

Db 361 gggatgt 420

Oy 1015 gaagcctatgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1074

Db 421 gaagcctatgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480

Oy 1075 agcagaatcagcctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1134

Db 481 agcagaatcagcctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 540

Oy 1135 aaagcctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1194

Db 541 aaagcctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 600

Oy 1195 ggcctcacaagcctcgtggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1254

Db 601 ggcctcacaagcctcgtggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 660

Oy 1255 tggcagcttttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1314

Db 661 tggcagcttttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720

Oy 1315 ttatagcaacaagagatccctcgtatgataataatgctgagcagcagcagcagcagcagcagc 1374

Db 721 ttatagcaacaagagatccctcgtatgataataatgctgagcagcagcagcagcagcagcagc 780

```

Qy 1375 cccaagtgacctcagcgtctcgttccctccacatggtctccagttcagcctaata 1434
    |||||||
Db 781 cccaagtgacctcagcgtctcgttccctccacatggtctccagttcagcctaata 840
    |||||||
Qy 1435 aatgaacccaatccagtgagcagcgaagaagaaacacaggtgtcgaagttgtgttc 1494
    |||||||
Db 841 aatgaacccaatccagtgagcagcgaagaagaaacacaggtgtcgaagttgtgttc 900
    |||||||
Qy 1495 agcagctgtcttgaagcagcagcagagtggttaacaggttccatcagctgtcttga 1554
    |||||||
Db 901 agcagctgtcttgaagcagcagcagagtggttaacaggttccatcagctgtcttga 960
    |||||||
Qy 1555 tgtgtccctgtgagcgttgagccttccccaagaagtgaacctaac 1602
    |||||||
Db 961 tgtgtccctgtgagcgttgagccttccccaagaagtgaacctaac 1008
    |||||||

RESULT 9
AADI7515
ID AADI7515 standard; DNA; 8191 BP.
XX
AC AADI7515;
XX
DE 10-DEC-2001 (first entry)
XX
XX Human taste receptor, htr1r1 full-length genomic DNA.
XX
XX Human: taste-cell-specific G protein-coupled receptor; T1R1; drug;
KW genetic modulation; pharmaceutical; taste sensation; food industry;
KM chemosensory transduction; ds.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..7771
XX FT /tag= a "Human taste receptor, htr1r1 protein"
XX FT /product= "Human taste receptor, htr1r1 protein"
XX FT 1..846
XX FT /tag= b
XX FT /number= 1
XX FT /cons_splice= (5'site:NO, 3'site:NO)
XX FT 847..1037
XX FT /tag= c
XX FT /number= 1
XX FT 1038..1593
XX FT /tag= d
XX FT /number= 2
XX FT 1594..1900
XX FT /tag= e
XX FT /number= 2
XX FT 1901..2819
XX FT /tag= f
XX FT /number= 3
XX FT 2820..3581
XX FT /tag= g
XX FT /number= 3
XX FT 3582..4603
XX FT /tag= h
XX FT /number= 4
XX FT 4604..4816
XX FT /tag= 1
XX FT /number= 4
XX FT 4817..5138
XX FT /tag= j
XX FT /number= 5
XX FT 5139..5259
XX FT /tag= k
XX FT /number= 5
XX FT 5260..6841
XX FT /tag= l
XX FT /number= 6
XX FT 6842..7771
XX FT /tag= m

```

```

FT /number= 6
FT Intron 7772..8191
FT /tag= n
FT /number= 7
FT /cons_splice= (5'site:NO, 3'site:NO)
PN WO200166563-A2.
XX
PD 13-SEP-2001.
XX
PF 07-MAR-2001; 2001WO-US07265.
XX
XX 07-MAR-2000; 2000US-0187546.
XX 07-APR-2000; 2000US-0195536.
XX 06-JUN-2000; 2000US-0209840.
XX 23-JUN-2000; 2000US-0214213.
XX 17-AUG-2000; 2000US-0226448.
XX 03-JAN-2001; 2001US-0259227.
XX
XX (SENO-) SENOMYX INC.
XX
XX Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
XX WPI, 2001-582267/65.
XX P-PSDB; AAE10372.
XX
XX New mammalian taste-cell-specific G protein-coupled receptor
XX polypeptides for identifying compounds that modulate taste signaling
XX are useful in food, to modulate the sweet taste of foods or drugs -
XX
XX Claim 1; Page 78-81; 119pp; English.
XX
XX The invention relates to mammalian taste-cell-specific G protein-coupled
XX receptors, T1R and their corresponding cDNA molecules. Taste receptors,
XX T1R are useful for screening compounds which are used to activate or
XX modulate chemosensory transduction, such as taste sensation. The
XX identification and isolation of novel taste receptors and taste
XX signalling molecules allow for new methods of chemical and genetic
XX modulation of taste transduction pathways. The taste modulating
XX compounds are useful in pharmaceuticals and food industries to improve
XX the taste of a variety of consumer products, or to block undesirable
XX tastes, e.g., in certain pharmaceuticals. T1R's are also useful in
XX biochemical assay for identifying tastant (T1R) ligands having binding
XX specificity for T1R involved in taste signalling. The present sequence is
XX human taste-cell-specific G protein-coupled receptor, htr1r1 full-length
XX genomic DNA.
XX
SQ Sequence 8191 BP; 1695 A; 2332 C; 2162 G; 1902 T; 100 other;

Query Match 36.0%; Score 908.6; DB 22; Length 8191;
Best Local Similarity 99.7%; Pred. No. 1.9e-219;
Matches 930; Conservative 1; Mismatches 0; Indels 2; Gaps 2;

Qy 1594 gacctacagatgcagcctgttggaagaagaagtgagcactgaaggaaagccagacc 1653
    |||||||
Db 6841 gacctacagatgcagcctgttggaagaagaagtgagcactgaaggaaagccagacc 6900
    |||||||
Qy 1654 tgcctccgcgcacgtgtgtttttgtgcttggtgagcaaacctcttggtgtcgtc 1713
    |||||||
Db 6901 tgcctccgcgcacgtgtgtttttgtgcttggtgagcaaacctcttggtgtcgtc 6960
    |||||||
Qy 1714 gtagctaaacgtgtgtctgtcgtcgtcttgagacgtcgtgctgttccctggcac 1773
    |||||||
Db 6961 gtagctaaacgtgtcgtcgtcgtcgtcttgagacgtcgtgctgttccctggcac 7019
    |||||||
Qy 1774 ctgagacccctgtgtgtgagtgtaagcaggggcccgcctgtgttcttaagtggctcc 1833
    |||||||
Db 7020 ctgagacccctgtgtgtgagtgtaagcaggggcccgcctgtgttcttaagtggctcc 7079
    |||||||
Qy 1834 ctggcagcagtagtgagcagcctctatggtctctttgggaaacccaagcctcgtgc 1893
    |||||||
Db 7080 ctggcagcagtagtgagcagcctctatggtctctttgggaaacccaagcctcgtgc 7139
    |||||||

```


OY	1894	ttgttaagccaaagccctcttgcccttggttttcaacattctctctcgtcctgaacatt	1953
Db	7140	ttgttaagccaaagccctcttgcccttggttttcaacattctctcgtcctgaacatt	71399
OY	1954	cgctcatctcaactcaatcatcatcttcaagattttcacccaagatgacattcatcacc	20133
Db	7200	cgctcatctcaactcaatcatcatcttcaagattttcacccaagatgacattcatcacc	72559
OY	2014	gctctgggttccaaaacacaggtgtctggtcgtgtttgtgtatcatagctcagggccagctg	2073
Db	7260	gctctgggttccaaaacacaggtgtctggtcgtgtttgtgtatcatagctcagggccagctg	7319
OY	2074	cttctcgtcttaacttggcgtgtgtgtgtgtgaacccacctgcctgcctgagggaatacacagcgc	21333
Db	7320	cttctcgtcttaacttggcgtgtgtgtgtgtgaacccacctgcctgcctgagggaatacacagcgc	7378
OY	2134	ttcccccacatctggltgatgtctgtatgtgacagagaaacaaatcccttggtcttcatactggcc	21933
Db	7379	ttcccccacatctggltgatgtctgtatgtgacagagaaacaaatcccttggtcttcatactggcc	7438
OY	2194	ttctctctacaagtgtcctctctcttcatacagtgtgccttgcctgtcagcttaccctggtaagac	22533
Db	7439	ttctctctacaagtgtcctctctcttcatacagtgtgccttgcctgtcagcttaccctggtaagac	7498
OY	2254	ttggcagaagaactcaacaacgagggccaaatggttcaaccttcagcctgctcttcaactctgtg	23133
Db	7499	ttggcagaagaactcaacaacgagggccaaatggttcaaccttcagcctgctcttcaactctgtg	7558
OY	2314	ttctctgatactgacctctcttcacaacacgagccagcgtctactaagacgaggaatctactgcttgccgc	23733
Db	7559	ttctctgatactgacctctctcttcacaacacgagccagcgtctactaagacgaggaatctactgcttgccgc	7618
OY	2374	aacctgtagtgccttggtcgtcgtgagacacgtcgtgagcagcggcttcgtgtggtattttctgcctaa	24333
Db	7619	aacctgtagtgccttggtcgtcgtgagacacgtcgtgagcagcggcttcgtgtggtattttctgcctaa	7678
OY	2434	ttgtcaactgtatccctctgcgcgcgcgcagacccccaagaagacagacacattccagctccatct	24933
Db	7679	ttgtcaactgtatccctctgcgcgcgcgcagacccccaagaagacagacacattccagctccatct	7738
OY	2494	caggactacacgagagcgctgcggtcgcctcacctga	2526
Db	7739	caggactacacgagagcgctgcggtcgcctcacctga	7771
RESULT 10			
AADI19504			
ID	AADI19504	standard; cDNA; 705 BP.	
XX	AADI19504;		
XX	18-DEC-2001 (first entry)		
XX	Human novel G-protein coupled receptor (NGPCR) cDNA #4.		
KW	Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;		
KW	antisense-therapy; signal transduction; behavioural disorder; obesity;		
KW	heartbeat rate; inflammation; immune disorder; diabetes; cancer;		
KW	coronary disease; ss.		
XX	Homo sapiens.		
XX	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	1..705		
FT	/*tag= a		
FT	/product= "Human NGPCR protein"		
PN	W0200172842-A2.		
XX	04-OCT-2001.		
XX	28-MAR-2001, 2001MO-US09996.		
PF			

XX	28 -MAR-2000; 2000US-192978P.
PR	(LEXI-) LEXICON GENETICS INC.
XX	
PA	Walke DW, Scoville J, Donoh G, Turner CA, Friedrich G, Abuln A;
PI	Zambrowicz B, Sands AT;
XX	
DR	WPI: 2001-616474/71.
DR	P-PSDB; ABEL1972.
XX	
PS	Disclosure; Page 78; 80pp; English.
CC	
CC	The present sequence is a cDNA encoding human novel G-protein coupled
CC	receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
CC	the cellular membrane and are involved in signal transduction after
CC	ligand binding. The NGPCR polynucleotide sequences, are useful
CC	in diagnosis and treatment of a disease involving NGPCR, for detecting
CC	mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC	disease, for screening drugs effective in treatment of symptomatic or
CC	phenotypic manifestations of perturbing the normal function of NGPCR in
CC	the body or abnormalities in the signal transduction pathway mediated by
CC	NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC	trial monitoring and/or the treatment of physiological (heartbeats rate)
CC	or behavioural disorders. NGPCR is useful for identifying compounds
CC	useful in the therapeutic treatment of obesity, inflammation, immune
CC	disorders, diabetes, heart and coronary disease, metabolic disorders and
CC	cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC	genomic library which is helpful for identifying polymorphisms,
CC	determining the genomic structure of a given locus/allele and designing
CC	diagnostic tests.
XX	
SQ	Sequence 705 BP; 130 A; 231 C; 170 G; 173 T; 1 other:
	Query Match 27.9%; Score 704.6; DB 22; Length 705;
	Best Local Similarity 100.0%; Pred. No. 2,8e-168;
	Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1822 atgctggagctccctcgtgcagcgagtatgtagcagctctatgaccttttggggaaaccaca 1881
Db	1 atgcctggagctccctcgtgcagcgagtatgtagcagcctctatgaccttttggggaaaccaca 60
OY	1882 aggcctgctgcttcctacgcagcgagcctctttgcccttgatttcaacatctctgtccc 1941
Db	61 aggcctgctgcttcctacgcagcgagcctctttgcccttgatttcaacatctctgtccc 120
OY	1942 tgcctgacagttgcgtcatctccaactaatcatcatcttaagaagttttccacaaggatact 2001
Db	121 tgcctgacagttgcgtcatctccaactaatcatcatcttaagaagttttccacaaggatact 180
OY	2002 acattctcacagcgtctgggttccaacaacaggttgttgccgtttgtatatagagcca 2061
Db	181 acattctcacagcgtctgggttccaacaacaggttgttgccgtttgtatatagagcca 240
OY	2062 ggcggccagcttgcttatctgtctaactcttgctgctggtggtgtgaaccccaactcgtctgtag 2121
Db	241 ggcggccagcttgcttatctgtctaactcttgctgctggtggtggtgaaccccaactcgtctgtag 300
OY	2122 gaataccagcgtctcccccatctggtgatgtcttgatggtgacacagagaccaactccctgggc 2181
Db	301 gaataccagcgtctcccccatctggtgatgtcttgatggtgacacagagaccaactccctgggc 360
OY	2182 ttcaatcagcgtctctctataatgagctccctccatcagtagcctttgctgctgagagctac 2241
Db	361 ttcaatcagcgtctctctataatgagctccctccatcagtagcctttgctgctgagagctac 420
OY	2242 ctgggttaagagcttgcagagaaactacaacagagcgcaaatggttacaccttaagcctgtgc 2301

Db 1463 atgtgtcctggtacaccccaacaacacggtccctgtctccatgtgtgtcccaagagctgac 1522
 Qy 1505 ttgaagagcacacagcagtggttacggtttccatcactgtgtgtgagtggtgacct 1564
 Db 1523 agccagggcacaatgaaagatctgtgtgctccacacctgtgtgtcgtgagtggtgatt 1582
 Qy 1565 gtggggctgtggacctctcccaacaagaagt--gacctctacagatgcccagctgtgtgga 1621
 Db 1583 gtagcgacagcactaccctcaaccgctcagcagatgatttaactgtctgtccgtcccg 1642
 Qy 1622 aagaagaagtggtgacctgtggagagcagcagcctgtcccgagcactgtgtgtttttg 1681
 Db 1643 gtccatgtgtgtctcaacaagaacatcatcattgtccagcggtgacctctctgtg 1702
 Qy 1682 ctltgtgtgtagcacacctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1741
 Db 1703 agtggcagcagatgt 1762
 Qy 1742 tgt 1801
 Db 1763 cactgt 1822
 Qy 1802 gggggcgcctgt 1861
 Db 1823 gt 1882
 Qy 1862 gctctctgt 1921
 Db 1883 tgtatgt 1942
 Qy 1922 gtttccatcttctgt 1981
 Db 1943 gcttctccatctgt 2002
 Qy 1982 agtttcccaacaggtacatctacatcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2041
 Db 2003 agatgt 2062
 Qy 2042 tgt 2098
 Db 2063 tctgt 2122
 Qy 2099 tgt 2158
 Db 2123 ccacatcaaccatctgt 2182
 Qy 2159 gcacagagacacactcctgt 2218
 Db 2183 gccaccttaactcgcgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2242
 Qy 2219 tcaagtctgt 2278
 Db 2243 tgt 2302
 Qy 2279 aatgt 2338
 Db 2303 agttctacactcagcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2362
 Qy 2339 ccaggtctacagcagcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2398
 Db 2363 tgt 2422
 Qy 2399 tgaagcagcgt 2458
 Db 2423 tctgt 2482
 Qy 2459 acctcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2510
 Db 2483 agcgcaacacactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2534

RESULT 15
 AA258964

ID AA258964 standard; DNA; 2532 BP.
 AC AA258964;
 XX
 DT 08-MAY-2000 (first entry)
 DE Mouse GPCR-B4 polypeptide encoding DNA.
 XX
 DE Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; mouse;
 XX taste transduction pathway; taste receptor; foliate; fungiform; food;
 KW circumvallate; taste signaling; pharmaceutical; ds.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2532
 FT /tag= a
 FT /product= "GPCR-B4"
 FT
 PN WO200006593-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 27-JUL-1999; 99WO-US171104.
 XX
 PR 28-JUL-1998; 98US-0095464.
 XX
 PR 17-DEC-1998; 98US-0112747.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Zuker CS, Adler JE, Lindemeier J;
 XX
 DR WPI: 2000-195257/17.
 DR P-PSDB: AAY77557.
 XX
 PT New isolated sensory transduction G-protein coupled receptor. useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway and for generating taste topographic maps -
 XX
 PS Claim 5; Page 71-72; 76pp; English.
 XX
 CC The invention provides nucleic acids encoding rat, mouse and human
 CC sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
 CC GPCR polypeptides are components of the taste transduction pathway. The
 CC nucleic acids can be used to identify taste cells and as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
 CC for identifying subpopulations of taste receptor cells such as foliate,
 CC fungiform, and circumvallate taste receptor cells. The polypeptides can
 CC be used for identifying compounds that modulate sensory signaling in
 CC sensory cells. Such modulators of taste transduction are useful for
 CC pharmacological and genetic modulation of taste signaling pathways. These
 CC modulatory compounds can then be used in the food and pharmaceutical
 CC industries to customize taste. The present sequence represents a DNA
 CC encoding a mouse GPCR-B4 polypeptide.
 CC
 CC Sequence 2532 BP; 498 A; 824 C; 631 G; 579 T; 0 other;
 SQ
 Query Match 16.6%; Score 420; DB 21; Length 2532;
 Best Local Similarity 50.6%; Pred. No. 4,9e-96;
 Matches 1241; Conservative 2; Mismatches 1167; Indels 42; Gaps 8;
 Qy 89 ctgactaacccctcccgagatctactctctgagcagcctgttccctccatctcgtgt 148
 Db 83 ccgacttaaccgt 142
 Qy 149 gtctgaggt 205
 Db 143 tgaagagtgctctcaccctcagctactcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 202
 Qy 206 atgagatgt 265

[illegible]

D	1274	tcagggagatctgcatgttcaacttcaactcagctctctcggcaaccaagctctcttcttcgacggac	1333
Q	1325	acagagatcccccctcagtagctataacaatatgctcgtggacttggaaatggaccaagtggaa	1384
D	1334	aaggggagacatgcgcatgtactctctcgacatcatccatcgaatgggcagtgggccttgagccagaacc	1393
Q	1385	cccttcaacggtctcgtgcttctctccacatagtctccaggttcacgtttaacacataatggagaca	1444
D	1394	cccttccaaagcactcgctctctactctcccccacaggaaagcaggtgcacacatactt---agca	1450
Q	1445	aaatccagtttggcagcaggaaagcgaacacacagcttggtcctaagtcgtgtgttccacgagcttc	1504
D	1451	atgtgtctcgtgtaacaccccccaacaacacagcttcccacatatcatgttcttaagagtctgc	1510
Q	1505	ttgaaaggcaccagcagagtgtgttaccggtttccatcatcgtctgtcttgaatgtgtgccct	1564
D	1511	agccctgggcaaaatgaaaaaaaacccataagagcttccacacacatgtctctcgaagtgtgtgact	1570
Q	1565	gtggggctggggaccttctctcaacaagaagt---gacctacagatggacagctctgttggga	1621
D	1571	gtccgcgcggagacccctacacttcaaccgacatcaagtatgaatgaaactgtctgtctccgcgcg	1630
Q	1622	aagaagagatgggcaccttggcagtggaagcagacatgtctccgcgcacatgttgggtttttg	1681
D	1631	gtccatctgtgtcttcaagaacaacacatcgtctgttccaagcggcgtgtgccttctcgtg	1690
Q	1682	ctttcgtgtgagcacacacctcttggtgtgtgtcgtcgtcgacagctcaaacgtctgtctgtcgtgc	1741
D	1691	agttgcacagaaagtgcgcacatcgtgtgtgtgacatccctcgtgcgcgccttgggtcttcaacagta	1750
Q	1742	tgcctgggagactcgtgcctgtttgtgcctcgtgcacactaagaaaccccttgggttggagtcaagag	1801
D	1751	cgcttgccatctctgttcaactctctcgtgagagacatttccacagaccacatgtgtgcctcggcg	1810
Q	1802	ggggcgcgcgtgtcttcttactgtcgtgcctccctcgtgcagcaggtagttgagcagctctatg	1861
D	1811	ggggccccaatgtgtctctctcgtatgtcgtgtgcctcgtcgtcgtcgttcgggttggtgtcccg	1870
Q	1862	gctctcttgggaaaccccaagaagcctcgtgtgtgtcgtacagcagggccctcttggccttg	1921
D	1871	tgtatgtgggcccccccaaggtctctctcgtgttctcgtcgcgcagagcttcttcaacgcttt	1930
Q	1922	gttcaacatctctcgtctcgtcgtcgtacagttcgtcatcttccaaactaatacatcatctca	1981
D	1931	gcttctcgtctgcctctccctctcgtacatcaacggtgctgtctcttccacagattgtgtgtgccttca	1990
Q	1982	agtttccacccaaggtactcatctatcttccacacagcctggttccaaacacaggtgtgtgcg	2041
D	1991	agatvgccaagacgctgcgccaagcgcctcagttcttcgtgagtgcgttaaccaagggccttaag	2050
Q	2042	tgtttgtatgtacagctcagctcagcggccagctg---cttatcgttcaactctgtgcgttgg	2098
D	2051	tcttggtgcttcaatcaacgycggtcagaagtvggcccgtgtgtgcgggcaaatgctgcgca	2110
Q	2099	tgttgaccaccaactgcgtcgttagggaaataccacagcgtctccccaactcgtgtgagtcttgaqt	2158
D	2111	ccacacataaaccccatgtgcgcgacccgagaccccgatgacccaatactataactctctct	2170
Q	2159	gcacagagacaaactccctcgtgcttcaactacgtgccttccatacaatgagcctccctcca	2218
D	2171	gcacaccctaactaccggaacgggtgaactcttcaacaacagcatggaacttgtctgtctgcg	2230
Q	2219	tcaagtgccttgcgtcagctacactcgtggtaaagacttgcgcagagaaatacaacagagagcca	2278
D	2231	tgtcgtgtttcagctctcgtctacgttgggcagaagaaactgccacaacatacaacagagacca	2290
Q	2279	aatgtgtacacttgaagcctcgtccttcaactctcgtcttgatagcgtctcttcaacaaagg	2338
D	2291	agttcatcaaccctcaggaatgactcttcccttaactcttcccaactctccctctgtgcaagttca	2350
Q	2339	ccaggtctatcagcagcaagtaactcgtcccttgcgcgcaaaatgatatgtcgtgtgtgagcagcc	2398
D	2351	tatctatcacacataagctctcgtatcgtacacatataatgatctctcgtgtgactctgtacact	2410

QY 2399 tgaagcagcggtctcgtgtggatatttctgcctaagtgtaogtgaatcctctgcccag 2458
Db 2411 ttctggccatcggtctgggtacttggcccaaatgttaacatgatcctttctaccgg 2470
QY 2459 acctcaacagacagagcacttccaggcctccattcagactacacagagcg 2510
Db 2471 agcgcaacacttcagcttattcattagcatgaltcagggtacacagatgag 2522

Search completed: June 11, 2002, 06:48:12
Job time: 10293 sec

[illegible]

QY 2101 tgaacccactgctctgtaaggaaataaccagcgtctcccccactctgctgtagctgagtc 2160
|||||
Db 2101 TGAACCCACTGCTCTGTAGGAAATACAGCGCTTCCCCCATCTGTGTATGCTTAGTGC 2160
QY 2161 acagagacaaactccctggtctcattactggtcctctcctacaatggtcctcctcactc 2220
Db 2161 ACAGAGAACCACTCCCTGGGCTTACTGTGCTCTCTCAATGAGGCTCTCTCCATC 2220
QY 2221 agtgccttctgctcagctactggtgtaaggaacttgccaggaactataaagagccaa 2280
|||||
Db 2221 AGTGCCTTGTGCTGAGTACTGTGGTAAGGACTTGGCCAGAGAACTACAAAGAGGCCAAA 2280
QY 2281 tgytcaacctcagcctctctcaactctgtctccctgagtcgacctcttaccagagcc 2340
Db 2281 TGYTCACACTTACAGCTCTCTTCACTTCGTCTGTGATCGCTTCTTTCACACAGGCC 2340
QY 2341 agcgtctacagagcaagtaactctgctgagccaaatgagtgtggtgagcaagcctg 2400
Db 2341 AGCCTCTACGACGCAAGTACTGCTGCGGCCAAACATGATGGCTGGAGCAGCCTG 2400
QY 2401 agcagggcttcggtggtattcttcgtctaagtgtctacgtagctctctgcccagac 2460
|||||
Db 2401 AGCAGCGGCTTGGTGGGTATTTCTGCTAAGTCTACGTGATCTCTGCGCCAGAC 2460
QY 2461 ctcaacacagagagcaactctcagagcctcactcaaggaactaacagagcgctgcgctcc 2520
Db 2461 CTCAACACAGACAGACTTCCAGGCCCTCCATTCAGAGACTACAGAGCGCTCGCGCTCC 2520
QY 2521 acctga 2526
Db 2521 ACCTGA 2526

RESULT 2
AX282900 2951 bp DNA linear PAT 02-NOV-2001
LOCUS AX282900
DEFINITION Sequence 9 from Patent WO0172842.
ACCESSION AX282900
VERSION AX282900.1 GI:16609872
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
Abuin,A., Zambrowicz,B. and Sands,A.T.
TITLE Novel human 7tm proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0172842-A 9 04-OCY-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source location/Qualifiers
1..2951
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 559 a 888 c 831 g 669 t 4 others
ORIGIN

Query Match 100.0%; Score 2524.8; DB 6; Length 2951;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 ctgttgaagagttctgttagcttcaatgagcatggtctacacactctccaaagtatgcg 240
Db 412 CTGTGTGACAGAGTCTTTGAGCTTCAATGACATGGCTACCACTTCTTCAGAGCTATGGCG 471
QY 241 ctgtggtcttgagagataaacaactccacagccctctgctcccaacatcaacctggtgtac 300
Db 472 CTGTGGGTGTGAGAGATTAACACACTCCACAGGCCCTGCTGCCAACATCACCTGGGGTAC 531
QY 301 cagctgtatgagtgtgtgtgtgactctgccaatgtgtatgccaagcttgagaagtctctcc 360
Db 532 CAGCTGTATGATGTGTGTGTGACTCTCCATGTGTGTGTCAGAGCTGAGAGTGTCTTCC 591
QY 361 ctgtccagaggaacaccacataagagcttccaagaagacctctccactatctccaaaggtg 420
Db 592 CTGCCAGGGCAACACACATGAGTCCAAAGAGACCTTCTCCACTATTTCCCTTACGGGTG 651
QY 421 ctgtgaagtattggtgctgacagacccaacagctgtctccacaacagccgacctgtgagc 480
Db 652 CTGGCAGTGTATTTGGGCTGTGACAGCACCAACCGTGTGTCACACACAGCCGCTGTGAGC 711
QY 481 ccttctctgtgtgccatgatagtatgctgagccagcagcagaagctcagctgtaagcgg 540
Db 712 CCTTTCCTGTGTGCCCATATTGATGCTATGCGCCAGCACGAGAGCCTCAGCGTGAACGG 771
QY 541 cagatatacctcttctctctgacacacatcccaatgataaagatccaggtgtgaaacatgtg 600
Db 772 CAGTATCCCTCTTCTCTGCGACATCCCAATGACAAAGTACAGGTGTGAACATCTGGTG 831
QY 601 ctgtctctgcaagaagttcgggtgtgacccctgatactctctgtctgttgagcagtgacacat 660
Db 832 CTGCTGCTGACAGAGTGTGGGTGTGACCTGTGATCTCTGTGTGGAGAGTGTGACACTAT 891
QY 661 gggcagctaaaggtgtgcaagcaactggaagaaaccaagccactgtgtcagggtatctgcaatgt 720
Db 892 GGGCAGCTAAGGGGTGACAGGCACTGTGGAACAGGCACTGTGTGAGGAGTGTGACTTGT 951
QY 721 ttcaagagcatcatagcccttctctgccaaggttgagagtgagatgacatgtctctatg 780
Db 952 TTCAAGGACATCATGCTCTTCTCTGCCCCAGGTGGGCGATGAGAGATCACTGCTCTCATG 1011
QY 781 cgcacactgtgcccagagccgggggcccacgctgctgtgtgttttttccagccggcagttgac 840
Db 1012 CGCACACTGTGGCGCAGAGCGCGGGGCCACGCTGTGTGTGTGTTCAGCGGCGAGTTGGCC 1071
QY 841 aggtgttttctgagctcgtgtgtgtgtgacccaactgtactgtgcaaggtgtgtgtgtgtc 900
Db 1072 AGGGTGTTTTTCAGAGTCCGTGTGTGTGACCAACTGACTGTGCAAGTGTGGGTGCCCTCA 1131
QY 901 gaaagctgtggccctctccagacacatactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Db 1132 GAAGCTGTGGCCCTCTCTCCAGGCATCTCACTGTGGGTGTGCCGGGATTCACAGCATTTGGATG 1191
QY 961 gtgtctgagcgtgtgcatccacaagaagaggtgttccctgtgagcgttgtaagaagcc 1020
Db 1192 GTGTGTGGCGGTGGCCATTCAGAGAGGGGTGTGCTGTGCTGTGAGAGGCTTTGAAGAACC 1251
QY 1021 tatgcccgggcagacaagaagagccctcagcctgtgccaagaggtcctgtgtgaagcagc 1080
Db 1252 TATGCCCCGGGCAAGACAMAGAGGCCCTTGTAGCCTTGTCCACAAGAGGGGTCTGTGCGCAGC 1311
QY 1081 aatacagctctcagaagaatgtcgaagcttcatgtracacacagatgccccaaactcaaac 1140
Db 1312 AATCAGCTCTGTGAGAAATGTCCAACTTTCATGTACACAGATGTCCAAAGCTCAAAAGCC 1371
QY 1141 ttctccatgagtctctgtccataacacgcataccgggtgtgtatgtgagtgccatgtgctc 1200
Db 1372 TTTCATGATGATTTGCTTACAAAGCATATCCGGGTGTGTATGGGTGTGCCATAGGCTTC 1431
QY 1201 caacagctctgtgagctgtgtcctctggaagctgttccaaagggcgagttacacctgtgag 1260
Db 1432 CACAGCTCTCTGTGGGT 1491

QY	1261	cttttggaagagatccaacaagtgagcttcccttcaacaagaagacgttgcttcat	1320
Db	1492	CTTTTGGAGCAGATCCACAAAGTGCAATTTCCTTCTTACAAAGGAACTGTGGCTTTAAT	1551
QY	1321	gacacagagatcccccctcagtagctaatacataatgccttgagctggatlgacccaag	1380
Db	1552	GACAAAGAGATCCCTCAGTACTATTAACATTAATGGCTGGAGCTGGAAATGAGACCAAG	1611
QY	1381	tgagacctcaaggtcctcggttccccaatggtctccagttcagctaaacataaayag	1440
Db	1612	TGGACCTTCACGGTCCGTGGTTCCTCCACATGGTCCAGTTCAACTAAACATAAATAG	1671
QY	1441	accaaaatccagtgagcaacggaagaagcaaacagtgagccaaagtcgtgtgttccagagac	1500
Db	1672	ACCAAAATCCAGTGGCAGCGGAAAGGACAAACAGATGCTTAATCTGTGTGTTCCAGGCAC	1731
QY	1501	tgctctgaaggagccaagcagagtggttaacggtttccatacactgtgctttgagtgtyg	1560
Db	1732	TGTCCTGAAGGCGCACAGCAGAGTGTTACGGGTTTCCATCACTGGCTTGAGTGTTG	1791
QY	1561	ccctgtgaggtctgagaccttccccaagaagtgaacctatacagatgccagcctgtgyg	1620
Db	1792	CCCTGTGGGGCTGGGACCTTCCCAACAAGAGTGAACCTTCACAGATGCCAGCTGTGTGG	1851
QY	1621	aaagaagaagtggagaaaccttgagggaagaacagagctgttcccgcgacgtggtgttttg	1680
Db	1852	AAAGAAAGTGGGACCTTAGGGAAGCCAGACCTGTCCCGGACGTGTGTGTTTGG	1911
QY	1681	gcttctggtgaagacacactcttggtgtctgtgagcagctaaacagctgtcgtgctgtg	1740
Db	1912	GCTTTCGCTGAGCACACTCTTGTTGGTGTGCTGTGGAGAGTAACAGCCTGTGTGTGCTG	1971
QY	1741	ctgcttgagagactgtcgctgttctgcctgagacatagacacccctgtgtgtgagtgacga	1800
Db	1972	CTGCTTGGGACCTGTGGCCCTGTTGGCTGAGCACTAAGACACCCCTGTGTGTGAGTGACCA	2031
QY	1801	ggaggagcagctgtgacttctatgctgagcgctccctcgagcagagtaagtggagacctcat	1860
Db	2032	GGGGGGCGGCTGTGTCTTTCTTAATGCTGGGTCCCTGTGGCAGAGTAAGTGCGCTCTAT	2091
QY	1861	gactctcttgagggaaccacaagaagcctgtgctgttgctacagcaagagcctcttgccctt	1920
Db	2092	GGCTTCTTGGGGAACCCACAAGGCGGTGGGTGTTGCTACGGCAGGCGCTTGTGGCTT	2151
QY	1921	ggtttcaacatctctctgtcttcgtcgtgacaagttgctcatcccaactaatcaatccttc	1980
Db	2152	GGTTTCAACACTTTCCTGCTCCCTCGACAGATGTGCTCAATTCACATATCATCTTTC	2211
QY	1981	aagttttccagcaagagactatacttccaaagcagctgtgtccaaacaacagagtgctgac	2040
Db	2212	AAGTTTTCACCAAGGACTATCAATTCACACAGCCTGGGTCCAAACACAGGTGCTGCG	2271
QY	2041	ctgttggtgaatcagactcagcagcgccagcctgactatcgtctaaacttgagctgtgtg	2100
Db	2272	CTGTGTTGTAAATGAACTAGCTAGGGGCGCAGCTGCTTATGTGTCAACTTGGCTGGTGGT	2331
QY	2101	tgagaccacatgctcgtctagggaaataacagcgcttccccaatcgtgtgatgctgtatgc	2160
Db	2332	TGGACCCCACTGCTGCTAGGGAATACAGAGCGCTTCCCACTGTGATGCTGATGAGTGC	2391
QY	2161	acagagaagccaactccctggtgctcatatcagcgtcctctctataatgtgectctctcatc	2220
Db	2392	ACAAGAGCAACTCCGTGGGCTTCATACGTGGGCTTCTCTACAAAGGCGCTCCTCCATTC	2451
QY	2221	agtgaccttgccgtcagcactacactggttgaaggactggccagagaacataaagaaggtcaaa	2280
Db	2452	AGTGCCCTTGTGCTTGACACTACTGTGGTAAAGACTTGGCAGAAACTACACGAGGCGAAA	2511
QY	2281	tggttcaaccttcagcagctgctcttcaactctgtgtccgtgatcgctcttcaacaagagcc	2340
Db	2512	TGTGTCACTTCAAGCGCTGCTTCAACTTTCGTGTCTGTGATGTGCTTCCACACAGGCGC	2571
QY	2341	agcgctcaagcgaagtaactcgtcctgcgccaacatgatygtcgtggcgtgagcagcctg	2400

Db	2572	AGCGCTACGAGGCAAGTACTGCTGGCGGCACATGATGCTGGGTGAGCAAGCTGC	2631
OY	2401	agcagcgcttcgtgtgtatcttcgtccaaagtgtactgtatctctgtccgccagac	2460
Db	2632	AGCAGCGGCTTCGTGGGGATTTCTTGCTAGTACTACGTATCTCTGCGCCGACAGC	2691
OY	2461	ctcaacagcaagagcactccagccctccatctcaaggactacaagagcgctgtggctcc	2520
Db	2652	CTCAACAGACAGACACTTCAGGCGCTTCATTCAAGACTACACAGAGCGCTGCGGCTCC	2751
OY	2521	acctga 2526	
Db	2752	ACCTGA 2757	
RESULT	3		
LOCUS	AX354027	2759 bp	DNA linear PAT 06-FEB-2002
DEFINITION	Sequence 13 from Patent WO0203846.		
ACCESSION	AX354027		
VERSION	AX354027.1	GI:18618964	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	Naito,A.T.		
JOURNAL	Method for opening the blood-brain barrier		
FEATURES	Patent: WO 0203848-A 13 17-JAN-2002;		
source	Naito, Albert T. (US)		
	location/Qualifiers		
	1..2759		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="Incyle ID No: 7475203CB1"		
BASE COUNT	523 a 828 c 782 g 626 t		
ORIGIN			
Query Match	99.9%;	Score 2523.2;	DB 6; Length 2759;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 2522;	Conservative 3;	Mismatches 1;	Indels 0; Gaps 0;
OY	1	atgcgcctctgcaagcgctcgccgtgtgcgcgcgcgcgcctcatcttcctgtcgtcggcc	60
Db	53	ATGCTGCTCTGCACGCGCTCGCTGCTGCGCGCTTCATCTTCTGCTGCTGCGCGC	112
OY	61	ttgtcctgcataagcaagagctctctcctgaactcaacccctccggagattactcctg	120
Db	113	TTTGCTGCGCATAGACAGGAGTCTTCTCCTGACTTACACCTCCCGGAGATTAACCTCG	172
OY	121	gcagcgctgttcctccctccatctctgtcgtctgcaggtgtgaagcaagaccgagttgac	180
Db	173	GCAGGCGCTTCCCTCCATCTCCATCTGCGTGTGACAGGTGAGGACACAGACCAGGTGACC	232
OY	181	ctgtgtgaacagctctgtgaacttaaatgacatgacatgacacacccctccaagctatgg	240
Db	233	CTGTGTGACAGGCTCTTGTAGCTTCAATGACATGCGCTACACACTCTTCCAGGCTATGCGG	292
OY	241	cttgggtgttgagagataaacaactcaacgacctgtctgtcccaacatcaccttgggttac	300
Db	293	CTTGGGTTTGAAGGAGATTAACAACATCCACAGGCGCTGTGCCCAATCACCCTGGGGTAC	352
OY	301	cagctgtatgatgtgtgttlytgactcttcgcaatgtgtatgcacagctgaagtgctctcc	360
Db	353	CAGCTGTATGATGTGTGTTCTGACTCTCCAAATGTTATGCCACGCTGAGATGCTTCC	412
OY	361	ctgcagagcaacacacatagagcttccaagagaccttccaactatccctcctcaggtt	420
Db	413	CTGCCAGGCGCAACCACTATGAGCTCCAGAGAGACCTTCTGCACATATTTCCCTACGGTG	472

QY 421 ctggcagtgatgtggccttgacagaccacacgctgtgcccacacagccgctctgagc 480
|||||
Db 473 CTGGCAGTGATGGAGCTTGACAGCACCAACGGTGTGCTGCCACCAAGCCGCTTGTGAGC 532
QY 481 ccttctctgtgtccatgatagctatgtcggccagcagcgagacgtcctcaagcgttaagcgg 540
|||||
Db 533 CCTTCCCTGGGTCACATGATAGCTATGCGGCAGCAGCGACGCTCAGCGTGAAGCGG 592
QY 541 cagatattcccttcttctctgagcaccatcccacatacaagtaacaggtacaggtgagacatgtgtg 600
|||||
Db 593 CAGATATCCCTTCTTCCGCGCACCATCCCAATGACAAGTACAGGTGGAACATGATGCTG 652
QY 601 ctgtctgtcagaaagttcaggtgtgagctgtgactctctctgtgtgtgcaagagtgagacact 660
|||||
Db 653 CTGTGCTGCGAAGATTGCGGTGAGACTGATCTCTCTGTGTGGCAGCAGTGAAGACTAT 712
QY 661 gggcagctaaaggtgtcagagcactgagagaacagagccactgtgtcaggggactgtcatgtct 720
|||||
Db 713 GGGCAGCTAGGGGTGCGAGGCACTGAGAACCCAGCCACTGTGTCAAGGGGATCTGATGCT 772
QY 721 ttcaagaaatcatatgccccttctctctcagcaggtgagcgatgagagatgacgtgctcatgt 780
|||||
Db 773 TTCAAGAACATCATGCCCTTCTCTGCCAGGTGGGCATGAGAGATGCAAGTGCCTCATG 832
QY 781 cgcacactgagccagagccggggccacgcgtcgtgtgtgttttctccagccggcagttgccc 840
|||||
Db 833 CGCACCTCGGCCAGGCCGCGGGCCACCGTGTGTGTTTTTTCAGCCGGCAGTTGGCC 892
QY 841 aggggtgttttctgagtcggtgtgtgtcagcaacactgagcctgtgcaaggtgtgtgtgtcgtcta 900
|||||
Db 893 AGGTTGTTTTTCAGTCCGTGCTGCTGACCAACTGACTGCAAGGTGTGTGGTGCSCCA 952
QY 901 gaaacccgtggccctctccagcacaatcactggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
|||||
Db 953 GAAACCTGTGGCCCTCTCTCCAGCACATCATGTGGGTGCCCGGAGTCCAGCGATTGGGATG 1012
QY 961 gtgtctgtgctgtgscatccagaagaagggcgtgtccctgtgscctgtgaagcgttttgaagaagcc 1020
|||||
Db 1013 GTGCTGGGCGGTGGCCATCCAGAAAGGGGCTGTCCCTGGCCGTGAAGGGGTTTGAAGAAGCC 1072
QY 1021 tatgtccgggagacaaaggaagggcccttaagcctgtgccaagggctcctgtgtcagcagc 1080
|||||
Db 1073 TATGCCCCGGGCGACAAAGAAAGCCCTAAGGCTTGCACAAAGGGGCTCTGTGTGAGCGC 1132
QY 1081 aatagcctctgcaagaagaatgtcagaacttctatgacacagatggtcccaagctcaaaagcc 1140
|||||
Db 1133 AATAGCTCTGCAAGAAATGCCAAGCTTTTCATGGCAGACAGATGCCCAAACTCAAAAGCC 1192
QY 1141 ttctccatgagttctgtcctaacaacgcataccgggctgtgtatgtcgtgtgtgtgtgtgtgtgt 1200
|||||
Db 1193 TTCTCATGAGTCTGTGCTTACAAACGATACCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1252
QY 1201 caccagctctctgt 1260
|||||
Db 1253 CACCAAGTCTCGGGCTGTGCTCTGTGAGCTTGTTCAGGGGCGCAGTCTACCCCTGGGAG 1312
QY 1261 ctttttgagcagatccacaaggtgacttcttcttcttcttcttcttcttcttcttcttcttctt 1320
|||||
Db 1313 CTTTGGAGCGATGCCAAGGTGATTTCTTCTTACAAAGAGACAGTGTGGCGTTTAT 1372
QY 1321 gacacagagataccctctcagtagctatacaataatgtcctgtgagctggaatgtgaccccaag 1380
|||||
Db 1373 GACAAAGAGATCCCTCAGTAGTATTAATATTTGGCTGGAGCTGGAATGGAAGCCCAAG 1432
QY 1381 tggaccttcaaggtctctgt 1440
|||||
Db 1433 TGGACCTTCAAGGTCCTCGGTTCTTCCACATGCTCTCCAGTTCAGCTAAACATTAATGAG 1492
QY 1441 accaaatccagtgagcaggaagaagacaaaggtgtcctaagctgtgtgtgtgtgtgtgtgtgtgt 1500
|||||
Db 1493 ACCAAATCCAGTGGACGGAAGGACAAACAGGTGCTTAAGTGTGTGTGTGTGTGTGTGTGTGTGT 1552
QY 1501 tgtcttgaagggcaccagcagtggttcaaggttccatcactgtcgtcttgaagtggtgtgt 1560

Db 1553 TGTCTTGAAGGGCACACAGGAGTGTACGGGTTTCCATTCATCTGCTGTTTGAATGTGTG 1612
QY 1561 cccctgtgggcttgggaccttctcacaagaagtgacctctacagatgtccagccttggg 1620
|||||
Db 1613 CCTGTGGGGCTGGGACCTTCTCAACAAAGATGACCTTCAATGATGCAACCTTGTGGG 1672
QY 1621 aaagaagaagtggaacctggaaggaagcagacccgtctccggcactgtgtgtgtgtgtgtgtgtgt 1680
|||||
Db 1673 AAAGAAGAGTGGACCTGAGGAAAGCCAGACCTGCTCCGCGACATGTTGTGTGTGTGTGTGTGTGT 1732
QY 1681 gctttgtgagcaaaccttctgt 1740
|||||
Db 1733 GCTTGGGTGAGCACACCTCTTGTGGGT 1792
QY 1741 ctgcttggagctgt 1800
|||||
Db 1793 CTGCTTGGAGCTGTGGCTGT 1852
QY 1801 gggggccgctgt 1860
|||||
Db 1853 GGGGGCCGCTGT 1912
QY 1861 ggtctcttggggaaccacaagccctgt 1920
|||||
Db 1913 GGCCTTGTGGGGAACCCACAAAGGCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1972
QY 1921 ggttccacatcttctgt 1980
|||||
Db 1973 GGTTCACCATCTTCTGT 2032
QY 1981 aagtttccacaaggtgaacctacatcttccagccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2040
|||||
Db 2033 AAGTTTCCACCAAGGATCTCATATTCACACGCTGTGGTCCAAACACAGGAGTGTGTGTGTGTGTGTGT 2092
QY 2041 ctgtttgtatgatcagctcagcggccagcgt 2100
|||||
Db 2093 CTGTTTGATGATGATCAGCTGACGCGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2152
QY 2101 tggaccccgactgctgtcaggaagatccagcgtctcccccactgtgtgtgtgtgtgtgtgtgtgtgt 2160
|||||
Db 2153 TGGACCCGACGCTGT 2212
QY 2161 acagagacaaatccctgt 2220
|||||
Db 2213 ACAAGACCACTCTCTGT 2272
QY 2221 agtgccttggcctgtcagctcagctgt 2280
|||||
Db 2273 AGTGCCTTGGCTGACGCTACCTGGGTAAAGGACTTGCAGAGAACTAACAGAGGCCAA 2332
QY 2281 tgytcaacctcaagcctgt 2340
|||||
Db 2333 TGTGTCACTTCACTGCTGT 2392
QY 2341 agcgtctacgagcaggaagtaacctgt 2400
|||||
Db 2393 AGGCTTACGACGGCAAGTACCTGT 2452
QY 2401 agcagcggcttgggt 2460
|||||
Db 2453 AGCAGCGGCTTGGGT 2512
QY 2461 ctcaacagcaagagcacttccagcctcactcaagagctacagaagcgtgtgtgtgtgtgtgtgtgt 2520
|||||
Db 2513 CTCAACAGCACAGACACTTCCAGGCCCTCCATTCAAGGACTTACAGAGCGCTGTGGGCTGCC 2572
QY 2521 accctga 2526
|||||
Db 2573 ACCTGA 2578

RESULT 4

AX282894
LOCUS AX282894 2292 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 3 from Patent WO0172842.
ACCESSION AX282894
VERSION AX282894.1 GI:16609869
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
Abidin,A., Zandrowicz,B. and Sands,A.T.
TITLE Novel human 7tm proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0172842-A 3 04-OCT-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source 1..2292
location/Qualifiers
BASE COUNT 437 a 690 c 638 g 524 t 3 others
ORIGIN
Query Match 90.7%; Score 2290.8; DB 6; Length 2292;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 235 atgcggcttgagggttagagataaacaactcacagcgccctctccacataccctg 294
DB 1 ATGCGGCTTGAGGAGTGAAGATAAACAACACTCACGCGCCCTGCGCAATCACCTG 60
QY 295 gggtaacagctglatagatgctgtgtgtgactctgcgaatgtglatgccagctgagatg 354
DB 61 GGGTACACAGCTGTATGATGTGTGTGTGACTCTGCGCAATGTATGCCAGCTGAGAGTG 120
QY 355 ctctccctgcagagggaacacacacataagctccaaaggacctcttcacattccct 414
DB 121 CTCTCCCTGCGCAGGGCAACACACATAGAGCTCAAGAGACCTTCTCCACTATTCCTCT 180
QY 415 acggtgtgcagagtagttagtgcctcacaacacacacacacacacacacacacacacac 474
DB 181 ACGGTGTGCGAGT 240
QY 475 ctgaagcccttctctgtgtgccaatgatactgacgagcagcagcagcagcagcagcagc 534
DB 241 CTGAGCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 535 aagcgagcagatccctctctctctgcgcacacatcccaatgacaagtacaggttgagacc 594
DB 301 AAGCGGCAATATCCCTCTCTGCGCACATCCCAATGACAACTGACAGGTGAGGACC 360
QY 595 atggtgtgcagctcagaaagctcgggtgagctgagatctctctctgttgagcagcagtgac 654
DB 361 ATGGTGTCTCTCTGCAAGAGTTCGGGTGAGACTGTGATCTCTGTGTGTGTGTGTGTGTGT 420
QY 655 gaactatggcagctagagggtgacagcactgagaaacacagcagcagcagcagcagcagc 714
DB 421 GACTATGGGCACTAGGGGTGACAGCACTGAGAAACACAGGCACTGTGTGTGTGTGTGTGT 480
QY 715 atgtgttcaagaacataatgcctctctctctgcgcaggtgtgagatgagagagatgacgtgc 774
DB 481 ATTTGCTTTCAAGACATCATGCTCTCTCTCTGCGCAAGTGTGTGTGTGTGTGTGTGTGT 540
QY 775 ctcatgagcagcagctgcccagagcgggacagcagcagcagcagcagcagcagcagcagc 834
DB 541 CTCATGCGGCACACTGGCCAGGCGCGGCGCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 835 ttggcagagggtgtttctcagatcagtgatgctgacacacacacacacacacacacacacac 894
DB 601 TTGGCAGAGGTTGT 660
QY 895 gctctcaagaagccttgagcctctccagagcagacatcactgaggtgctccggagatccagcgact 954

DB 661 GCCTGAGAGCCCTGGGCGCTCTTCACAGGCATCTGCGGCTCCGGGTCCAGCGCATTT 720
QY 955 gggatgtgtcgtggcgtgtgacatccagaagaagagctgtccctgcgcctgaagcgcttgaa 1014
DB 721 GGGATGTGTCTGGCGCTGGCCATCCAGAAAGAGGCTGTCTCCCTGAGCGCTTTTGA 780
QY 1015 gaagcctatgcccgggacagacaagagagcccttaggccttgccacaaggctcctgtgtgc 1074
DB 781 GAAGCTTATGCCCGGGACAGACAAAGAGGCGCTTAGCGCTTGGCAAGAGGCTCTGTGTGC 840
QY 1075 agcagaacatcagctctgcagagaatgcccagcttccatgacacacagatgcccagctc 1134
DB 841 AACACCAATACACTCTGCGAGAAATGCCAAGTTTCATGRCACACACGATGCCAAGCTC 900
QY 1135 aaagcctctccatgattctcctcacaagcatalaccggcgctgtatgcgttgccat 1194
DB 901 AAGCCTTCTCATGATGATTTCTGCTTCAACAGCATACCGGGCTGTGTATGCGTGGCCAT 960
QY 1195 ggcctccacagctcctctggcgtgtgctctgagctgtgtccaggggcccagctacccc 1254
DB 961 GGCCTCCACAGCTCCTGTGGGCTGTGGCTGTGGAGCTGTTCAGAGGGCCGAGTCAACCC 1020
QY 1255 tggcagctttggagcagatcccaaggtgcattccctctacacaaaggacgtgtggc 1314
DB 1021 TGGCAGCTTTTGGACAGATCCACAAAGGTGATTCCTTCAACAAAGACACTGTGGCG 1080
QY 1315 ttaatgacaacagagatccctccagtagctataacataattgctgtgagcgtgagatgga 1374
DB 1081 TTTAATGACAAAGAGATCCCTCACTAGCTATTAACATATTGCTCGGACCTGGAATGA 1140
QY 1375 cccaagtgcactcaagctcctcgttccacatgctcagatgctcaagttacgataacata 1434
DB 1141 CCCAAGTGACCTTCAAGGCTCTCGGTTCTCCACATGCTTCCAGTTCCAGTTAAACATA 1200
QY 1435 aatgagaccaaaatccagtgtgcagcggaaagacaaccaggtgtccaaagtctgtgtgtcc 1494
DB 1201 AATGAGACCAAAATCCAGTGGACGAAAGAACCAACAGGTGCTTAAGCTGTGTGTTC 1260
QY 1495 agcagctgtctgaaaggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1554
DB 1261 ACGAGCTGCTTGAAGGCGACCGACGAGGTGTACGGTTTCCATCAGCTGCTTTGAG 1320
QY 1555 tgtgtgccctgtgaggcctcctcacaagaagtgcctcacaagatgacatgacagcct 1614
DB 1321 TGTGTGCCCTGTGGGCTGGGACTTCCCAAGAGAGTACTCTACAGATCCAGAGCT 1380
QY 1615 tgtggaaagaagtgtggcactgtgaggaagcagcagcagcagcagcagcagcagcagc 1674
DB 1381 TGTGGAAAGAGAGTGGGACCTGAGGAAAGCCAGACTGCTTCCCGGCACTGTGTGT 1440
QY 1675 ttttggcttctgctgagacacacctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1734
DB 1441 TTTTGGCTTCTGCTGAGACACACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
QY 1735 ctgt 1794
DB 1501 CTGT 1560
QY 1795 tcaagcaggggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1854
DB 1561 TCAAGCAGGGGGCGGCTGT 1620
QY 1855 ctctatgtcttcttggggaacccaagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1914
DB 1621 CTCTATGCTTCTTGGGGAACCAAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
QY 1915 gcccttggttccacatctcctgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1974
DB 1681 GCCCTGT 1740
QY 1975 atcttcaagtttccacaaagttacttacttccacagcagcagcagcagcagcagcagcagc 2034

Db 1741 ATCTTCAAGTTTTCACCAAGTACTACATTTACACAGCGCTGGTCCAAACACACGCT 1800
Qy 2035 gctgacctgtttgatgatagctcagcgccagctgctatctgttacttgcctg 2034
Db 1801 GCTGGCCCTGTTGGATGATACGCTACAGGCGCCACGCTGATCTGCTACTGTGGCTG 1860
Qy 2095 gtgtgtgagacccacacgctcgtcaggaataacacagcgtcccccacatctgtatgctt 2154
Db 1861 GTGGTGGAGACCCACACGCGCTGCTAGGGAATACACAGCGCTTCCCCCATCTGTGATGCTT 1920
Qy 2155 gagtgcacagagacacactccctggtgcttactgtgctctctctacatgacctctc 2214
Db 1921 GAGTGCACAGAGACCAACTCCCTGGGCTTACTGTGGCTTCTCTACAAAGGCTCTCTC 1980
Qy 2215 tccatagtgacctgtgctgagctagctggttaagactgtccaggaataacacag 2274
Db 1981 TCCATGATGCTTGGCTTGGCTGAGCTACCTGGGTAAAGACTTGCACAGAACTACACAGAG 2040
Qy 2275 gccaaatggtcacccttcagctcgtctcctcactcgtctcgtcgtgctgctctcacc 2334
Db 2041 GCCAAATGTCACCTTACGCTGCTGCTCTCAACTTCGTCCTGTGATGCGCTTCTTCAAC 2100
Qy 2335 acggcaccagctctacagcgaagtaacctgctcgtcgcccaacatgagtgtggtgctgagc 2394
Db 2101 AGGCGCACGCTCTACGACGCGCAAGTACCTGCTGCGGCCAACATGATGCTGGGCTGAGC 2160
Qy 2395 agcctgagcagcgtctcgtggtggttatttctgctaaagtctcgtatcctctcgtcggc 2454
Db 2161 AGCCTGACACAGCGCTTGGGTGGGTATTTCTGCTTAAAGTCTAGTATCTCTGCGCGC 2220
Qy 2455 ccagacacctcaacacagacagacttccagagctccacatccacagacacagcagcgtcgc 2514
Db 2221 CCAGACCTCAACAGACAGACGACTTCCAGGCGCTCCATTCAGAGATACACAGAGGCGTGC 2280
Qy 2515 ggtccacactga 2526
Db 2281 GGCCTCACCTGA 2292

RESULT 5
LOCUS AF301162 2526 bp mRNA linear ROD 12-FEB-2001
DEFINITION Mus musculus strain 129P3/J putative sweet taste receptor T1R1
(Gpr70) mRNA, partial cds.
ACCESSION AF301162
VERSION AF301162.1 GI:12745519
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2526)
AUTHORS Li,X., Inoue,M., Reed,D.R., Hugue,T., Puchalski,R.B., Tordoff,M.G., Niinomi,Y., Beauchamp,G.K. and Bachmanov,A.A.
TITLE High-resolution genetic mapping of the saccharin preference locus (sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to mouse distal Chromosome 4
JOURNAL Mamm. Genome 12 (1), 13-16 (2001)
MEDLINE 21030739
REFERENCE 2 (bases 1 to 2526)
AUTHORS Li,X., Reed,D.R., Hugue,T., Puchalski,R.B., Tordoff,M.G., Beauchamp,G.K. and Bachmanov,A.A.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104-3308, USA
FEATURES
source Location/Qualifiers
1..2526
/organism="Mus musculus"
/strain="129P3/J"
/db_xref="taxon:10090"
/chromosome="4"
/map="Between D4Mit42 and D4Mit254"
1..>2526
gene

CDS
/gene="Gpr70"
1..>2526
/gene="Gpr70"
/codon_start=1
/product="putative sweet taste receptor T1R1"
/protein_id="AAK07092.1"
/db_xref="GI:12745520"
/translation="MLPFAHLLLSLQLAAYCAWAFSCORTESPSPSLPDELGLAGL
FSLHADLQVRRHPIVMSGDRSDSPSHGVHLROAMFPTVEELINSTRALLPNTTIGE
LHYDCESSNMYATIRYLAQGGHGLEMDQLNRHSSKVYALLGPDNTDHAFTTALL
SPLMLPLVSTEASSVILSGKRKFPSELTIPSDKYDEVIVRLQSGRWVILVGSY
GDYGOLOVALLELATPFGICVAFKDVPLVPSAAGDRMRQMLRLARATTVVVS
NRHLGAFPRSVVLANLTGKWLASEDMAISTYITNPGIOIGTVGVAVIQORVPS
KEEESYVOAVTGAPRTCEGSGNOCRECHAFETNMPELGFASAAVNYE
AYAVAHGLHQLIGTSRGCARGPVYPMOOLQIKYNFLHKKTYAFDNDGPIGY
DIYAMDNGEPMTEVYIGSASLSVPHDIDIKRTIQMHKRNQVPSVSTEDCLBGRH
LVMSHHCFCFCEHCEAGTFLNLSLHTCQPCGDEBAPGSSCSRYVEFLGMBP
ISLVLANTLILLLLIGTGLPAMRHTLVNVSAGRLCFMLGSLVAGCSLSYSE
GEPTVPACILROPFLSGLFAIFLSCLTIRSFOLVIFKEPSTKPTFHTMAOHGAI
FVIVSTVHLECLTFLAMNTPRETREFORPHVILECTEVSNGVLVAFAHILLS
ISTFVCSYLKELPENYNAKCVTFESLLHFVSLIAFTMSISYGSYLPVAVNLVAGL
ATLSGSGSYFLPCRYIILCRPELNTNHEHQASIDYTRRCGT"

BASE COUNT 483 a 760 c 679 g 604 t
ORIGIN

Query Match 67.0% Score 1692.8; DB 10; Length 2526;
Best Local Similarity 79.7%; Pred. No. 0;
Matches 1994; Conservative 3; Mismatches 505; Indels 0; Gaps 0;

Qy 22 ctgttcgagcagcagcttcattcctcgtcgtcgtggtccttgcctcagatagcagag 81
Db 25 CTGTCACAGCTGACAGCTGCGGTGCTTACTGCTGAGGCTTTCAGGTGCCAAGAGACAGAA 84
Qy 82 tctctcgtcagctcaccctcccgagagatactcctctgcaagcgtcttccctccat 141
Db 85 TCCCTCCAGGTTTCAGGCTCCGCTGGGACTCTCCCTGCGGACGGCTGTTCCCTCAT 144
Qy 142 tctgagctcgtcagagtgtaggcacagaccgagtgacccgtgtgacaggtctgtgagc 201
Db 145 GCTGACTTCTGACAGGTAGACACAGACCTCTGGTGACAAAGTTGTGACAGGTCTGACAGC 204
Qy 202 ttcaatgacatagcctacacaccttccagagctagcgttggtggttgtagagataaac 261
Db 205 TTCAACGGCCANGGCTATCACCTCTTCCAGACGANTGGGTTACACGTTGAGAGATAAAC 264
Qy 262 aactccacgcccgtcgtcccaacaatacacccttggtggtacacagctgtagtgtgtt 321
Db 265 AACTCCACAGCTCTGCTTCCCAACATCACCTGGGGTATGAACGTGTATGAGCTGTGCTCA 324
Qy 322 gactctgcacatgtagtagcagctgtagagtgctctcctccgacagggcaacacacata 381
Db 325 GAGTCTTCCATATGTTATGACACCTTGAGGGTGTCTCGCCGACGAAGGACAGCCACTTA 384
Qy 382 gagctcagaagacaccttcacatcctccctacacgtgctcgtgacagtagtctggcctgac 441
Db 385 GAGATGACAGAGATCTTCCGACACACACCTCCCAAGGAGTGGGACCATATTGGGCTGAT 444
Qy 442 agcaccaacgctgtagcacaacagcgccctgtagcgttccctcgtggtggtccatgatt 501
Db 445 AACACTGACACGCGTGTCCACCACTGCTGCTGAGACCCCTTTCATGATCCCTTGCTC 504
Qy 502 agctatgagcagcagagagagcgtcagctgtagagggagagtagtaccctcttccctgagc 561
Db 505 AGCTACGAGGCAAGAGGCTATCTCAGTGGAGAGGCCAAGTTCCCTCTTCTTGCGC 564
Qy 562 accatcccaatgacaagtacacagtgtagagacacatgctgctgctcagaaatctggc 621
Db 565 AOCATCCCGACAGATAATACAGGTGAGATGATAGTGGGCTGCTCAGAGGCTTCCGC 624
Qy 622 tggactgtagctctcgtgttgagcagagtagcagactatggcagagtaggggtgtagcga 681
Db 625 TGGCTGTGATCTGCTGTTGGCAGCTATGTTGACTACGGGACAGCTGGGCTACAGCGC 684

QY 682 ctggaacaccagccactggtcagggatctgattgcttcaaggacatcatccctc 741
 Db 685 CTGAGAGAGCTGGCAOCCACAGGGGATCTGGCTTCAAGAGCGTGGCTCTC 744
 QY 742 tctgccaaagtgtgagatgagatgacatgcccacccctggccagccgg 801
 Db 745 TCCGCCAAGCGGGGATCCCAAGATGACGCGATGATGCTGCTGCTCAGCAG 804
 QY 802 gcaacgctggtgtgtttttccagcagcagttggcaggtgttttcagtcctg 861
 Db 805 ACCACCGTGTGCTGGTCTTCTTAACCGGACCTGGCTGGATGTTCTTCAAGTCTG 864
 QY 862 gtgtgaccaaactgactgcaagagtgtgtgctcagaagccttgagccctcag 921
 Db 865 GTGCTGGCAACTGACTGGCAAGGTGTGATGCGCTCCGAAACATGGGCACTCCACG 924
 QY 922 caataactggtgtgtcccggtatccagcagcattggtgtgtgtgtgtgtgtgt 981
 Db 925 TACATCACCAATGTGCGCGGATCCAGGGATTTGGACGGTGTGGGGGTGGCCATCCAG 984
 QY 982 aagagggcgtccctgctgcaagagcgtttgaagaagcctatgcccggagagaag 1041
 Db 985 CAGAGACAGTCCCTGGCTTGAAGAGTTTGAAGAGTCTTATCCAGGCACTGACGGGT 1044
 QY 1042 gcccttaggccttgcacaaagggctcctgtgtgagcagaacatcagcttgcagaatgc 1101
 Db 1045 GCTCCAGAACTTGCCCAAGAGGGGTCTGTGGCGCACTAACACAGCTGTCCGGAGTGT 1104
 QY 1102 caagcttatgtacacacacagatgcccagctcaagccttccatgtgtgtgtgtgt 1161
 Db 1105 CAGGCTTTACGACATGAAACATGCCCGAGCTTGGAGCTTCTCCATGAGCGGTGCTTAC 1164
 QY 1162 aacgcataccagcgtgtgtatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1221
 Db 1165 AATGTGTAGAGAGCTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1224
 QY 1222 tctgtgagctgtgtcagagggcgagctctacccctgtgagccttggagagaatccaa 1281
 Db 1225 TCTGGGACCTGTGGCCAGAGGCCAGCTTACCCCTGTGACACTTCTTACAGAGATCTCAAG 1284
 QY 1282 gtgcatcttctctacacagagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1341
 Db 1285 GTGAATTTCTTCTTACATTAAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1344
 QY 1342 agctataacataatgtcctgtgagactgtgaatgagccaaagtgtgaccttccagctcctgt 1401
 Db 1345 TACTATGACATCATGCTGTGGAGCTGGAAAGGGCTGTGAATGAGACCTTTGAGGTGAT 1404
 QY 1402 tctccacatgtgtctcaatgtcagctaaacataatgagacaaatccaggtgacagga 1461
 Db 1405 TCTGCTCACTGTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1464
 QY 1462 aaggaacacaggtgtcctgaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1521
 Db 1465 AAGAAACATCAGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1524
 QY 1522 gtgtgtgaggtgtgtcctacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1581
 Db 1525 TTGGTATGGTGTCCACACAGCTGCTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1584
 QY 1582 ctcaaaagaagtgagacttcaagatgcccagccttgtgtgtgtgtgtgtgtgtgtgtgt 1641
 Db 1585 CTCACACACAGTGAAGCTTCAACAGCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1644
 QY 1642 ggaagcagcagcgt 1701
 Db 1645 GGGAGCTCAGCTGTCTTCTCAGCGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1704
 QY 1702 tgggt 1761
 Db 1705 TTGGTGTCTATTAGCACTTAACAGCGTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1764

QY 1762 ttgacctggcaccctagacacacccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1821
 Db 1765 TTGTGCTGGGCTTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1824
 QY 1822 atgt 1881
 Db 1825 ATGCTGGTTCCTGT 1884
 QY 1882 aggcctgt 1941
 Db 1885 GTACCGCGCTGT 1944
 QY 1942 tgcctgacagctgt 2001
 Db 1945 TGTCTGACATCGCTGT 2004
 QY 2002 acatttaccagcctgt 2061
 Db 2005 ACATTCTACACACTTGT 2064
 QY 2062 gggccagcgt 2121
 Db 2065 ACGGTCCATTGT 2124
 QY 2122 gaataccagcgt 2181
 Db 2125 GATGTCCATGCTGT 2184
 QY 2182 ttcatatgt 2241
 Db 2185 TTCTGT 2244
 QY 2242 ctggaagagcctgt 2301
 Db 2245 CTGGGTAAAGAACTGT 2304
 QY 2302 ttcaactgt 2361
 Db 2305 CTCACATCTGT 2364
 QY 2362 ctgacctgt 2421
 Db 2365 CTACCGCGGT 2424
 QY 2422 ttctgtcctgaagctgt 2481
 Db 2425 TTCCCTCCTAAATGT 2484
 QY 2482 cagacctcctatccagagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2523
 Db 2485 CAGGCTTCATCCAGAGCTTAACAGAGGCGCTGCGGCACTAAC 2526
 RESULT 6
 AF337040 2892 bp mRNA linear ROD 25-APR-2001
 LOCUS AF337040 Mus musculus candidate taste receptor T1R1 (T1R1) mRNA, complete cds.
 DEFINITION AF337040
 ACCESSION AF337040
 VERSION AF337040.1 GI:13785656
 KEYWORDS
 ORGANISM house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 AUTHORS Montmayeur,J.P., Liberles,S.D., Matsunami,H. and Buck,L.B.
 TITLE A candidate taste receptor gene near a sweet taste locus
 JOURNAL Nat. Neurosci. 4 (5), 492-498 (2001)
 PUBMED 11319557
 REFERENCE 2 (bases 1 to 2892)
 MONTMAYEUR,J.-P., LIBERLES,S.D., MATSUNAMI,H. and BUCK,L.B.
 AUTHORS
 TITLE Direct Submission

JOURNAL Submitted (18-JAN-2001) Neurobiology, Harvard Medical School, 220
Longwood Avenue, Boston, MA 02115, USA

FEATURES
Location/Qualifiers
source
1. .2892
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="4"
1. .2892
/gene="Trlr1"
183. .2711
/gene="Trlr1"
/note="G protein coupled receptor"
/codon_start=1
/product="candidate taste receptor Trlr1"
/protein_id="AAK39437.1"
/db_xref="GI:13785657"
/translation="MLEMAHLLLSIQLAAYAWCFSCQRTSSPFSPLGDFLLACL
FSLHADCLQVRRPLVYSCDRSDSFNGHGLFOAMFTEIINNSALLPNTLIGE
LYDVCSSESNVATLRYLAQOGTGLEMODLRNHSKVVALLGPDTHAVTAA
LSPRLMPLVSEASSVILSGKRKFPSEFRTIPSDKOYEIVRLLQSGMWISLGSY
GDYGOICVOALEELATPRGICVAFKDVPILSAODGPRMRMLRLARATYVYVS
NRHLAGFRSVVLANLTGKVMILASEDMALSTYITNPGIOGIGVLGVALIQORVPG
LKEPESYVQAVMGAPRTCEGSMCGNOLRCRCHAFETNMDELGFASMAVNYE
AVYAVAGHLQLIGCTSGTCARGPVYPMQLQOIYKNFLHKKVAFDQKPLGY
DIAMDNCEPEMTEFEVIGSASLSPVHLIDIKTKIOWHGKNNOPVSVCTDCEGHR
LVMSHHCECEMCEAGTFLNSELHOCPCGTEEMAPGSSACSEFVTEPFLGHR
ISLVLAANTLLLLILGTLGFLFAMRLHTVYVVSAGRLCFMLGSLVASCSISXSF
GKTPFACLRLOPLSLGFAIFLUSCLTIRSFOLVITFKESTKVPTEHTWAQNGAI
FVIVSVHLFLCLTWMAMTPTREYORPHLVITKESTKVPTEHTWAQNGAI
ISFVCSYLKELPENYENAKVTFSLHFLHVSIMIAFEFTSSITQGSYLPAVNVLAGL
ATLSGSGSFLPKCYIILCRPELNTEHFOASIQDYTRRCGTT"

BASE COUNT 585 a 841 c 794 g 672 t

ORIGIN

Query Match 66.9%; Score 1691; DB 10; Length 2892;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 1994; Conservative 3; Mismatches 508; Indels 0; Gaps 0;

QY 22 ctgttcgagctgagcttcaattcctctgctgctgagccttgcctgcataagacgag 81
DB 207 CTGCTCAGCCCTGCAGCTGGCGGTGCTTACTGCTGCGCTTACAGTGCACCAAGGACAGAA 266
QY 82 tctctctactaccacctcccgagatlaactcctcggagagcctggttccctcctcat 141
DB 267 TCTCTCCAGGTTTCAGCCTCCCTGGGAGCTTCTCTGGGAGGCGCTTCTCCCTCAT 326
QY 142 tctgagctgctgagctgagacacagccgagtgaccctgctgagagtgctgtgagc 201
DB 327 GCTACTGCTGCGAGGTGAGACACAGACTCTGCTGAGCAAGTGTGACAGGTCTGACAGC 386
QY 202 ttcaatgagcatggtacacacctcttccaggtatcgagcttggggttgatgagagataaac 261
DB 387 TTCAAGGCGCATGCTGCTATCACTCTTCCAAAGCATGCGGTTACACGTTGAGAGAGATAAAC 446
QY 262 aaccacagcgccctgctgcccacaatacaacctgggtgtacacgctgtagatgctgtgt 321
DB 447 AACTCCAGAGCTGCTGCTCCAAACATCACTCGGGGATATGAACTATGAGATGTGCTCA 506
QY 322 gacctcccaatgtgtatgacagctgagagtgctctccctgcagagggcaacacacata 381
DB 507 GAGCTTTCATATGCTATGCTATGCCACCTTGAGGCTGCTCGCCAGCAAGGACAGGCACCTTA 566
QY 382 gagctccaaggaacacctctccactatccctcaacgctgctgagagtgatggcctgagc 441
DB 567 GAGATGAGAGAGATCTTTCGCAACCACTCTCCAAAGGTGGTGCAACATATGGGCGCTAT 626
QY 442 agacccaacagctgtgctgcaacacagccgctctgctgagcccttccctggtgctccatgatt 501
DB 627 AACACTGACACGCTGTCCACACTGCTGCTGAGACCCCTTTTGTATGATCCCTTGCTC 686
QY 502 agcatatcgagcagacgagacgctcagctgtagcgagcatatccctcttccctgagc 561

DB 687 AGCTATGAGCGAGACAGGCTGATCTCAGTGGGAAGGCCAGATTCCCTCTTCTTGGCC 746
QY 562 accatcccaatgacaaatgaccaggttggaacacatgctgctgctgcagaaatctggg 621
DB 747 ACCATCCCGACGATTAATACAGGTGGAATGATAGGCGGCTCTCCACAGCTTCGGC 806
QY 622 tggacctgagctctctgcttgagcagctgagactatggcagctgaggggtgagagca 681
DB 807 TGGGCTTGATCTGCTGCTGTTGGCAGCTATGCTATGCTACGGGCACTGGGCTACAGCG 866
QY 682 ctggagaacacagcagctgctgagcaggtgagctgcatctgcttccaaggaacatcatgacctc 741
DB 867 CTGAGAGAGCTGGCCACTCCACAGGGGACATGCGTGGCTTCAAGAGAGTGGTGGCTCTC 926
QY 742 tctgccaagctgagcagatgtagagatgtagctgctcatatgctgcccactggtcccagcgag 801
DB 927 TCCGCGCCAGGCGGGTACCCCAAGATGCGAGGCAATGATGCTGCGCTCGAGCCAGG 986
QY 802 gcaacagctgctgctgcttcttccagccgagctgagcaggggtgttcttgaagtcggtg 861
DB 967 ACCACCGTGTGCTGCTGCTTCTCTAACCGGCACTGCTGGAAGTCTTTCAGGCTCTGTG 1046
QY 862 gtgctgacaaacctgactgagcaggtgtggtgcgctcagaagagcctgggacctcaccag 921
DB 1047 GTGCTGGCCCACTGACTGAGCAAGTGTGATCCGCTCCGAAGACTGGCCATCTCCACG 1106
QY 922 caacatcactggtgtgctccgagatccagcgcatgtggatgtgctgctggtgcgtgcatccag 981
DB 1107 TACATTCACCAATATGTCGCCGATCCAGGCAATGGGAGCGTGGGGGTGGCCATTCAG 1166
QY 982 aagagagctgctccctgagcctgaagagcgttgaagaagcctatgcccggagcagacaagag 1041
DB 1167 CAGAGACAACTCCCTGGCTTAAGAGAGTGTAAAGATCTTATGTCAGAGCACTGATGGGT 1226
QY 1042 gcccttgaagccttgcacaaagagctcctgctgctgagcagcaatcagctctgcagaaatgctc 1101
DB 1227 GCTCCCAAGAACTTGGCCAGAGGGGCTCTGTGGGCACTAACAGAGCTGACAGGAGTGT 1286
QY 1102 caagcttcatatgacacacagatgcccaagctcaagcctctccatgagttctgctaac 1161
DB 1287 CACGCTTTCAGACATGAGAACATGCCAGCTTGGAGCTTCTCATAGGCGCTGCCATTC 1346
QY 1162 aagcatacgggctgtgtatgctgctgagcgttgagccatgagcccaacagctcctggtgctgctc 1221
DB 1347 AATGCTATGAGGCTGTATGCTGTGCGCCACAGGCTCCACAGCACTCTGGAGTATCC 1406
QY 1222 tctgagagctgttccagagggcagagcttaccacctgagcagcttctggagcagatccacaag 1281
DB 1407 TCTGGAGACCTTGGCAGAGGCGCCAGTCTACCCCTGGCAGCTTCTTCAACAGATCTACAG 1466
QY 1282 gtgaattcctcttacaacaagaacatgctgagcgtttaaatagaacaagaatccctcagct 1341
DB 1467 GTGAATTCCTTCTTACATAAAGACCTGTACATTCGATTCGAAGGGGAGCCCTCTAGGT 1526
QY 1342 agcatatacaatattgctgagctgagctgagatgagcccaagctgagcttcaagtgctcgtgt 1401
DB 1527 TATTATGACATCATGCGCTGAGCTGAGTGAATGACCTGATGAGCTTGTGAGGTATTGGT 1586
QY 1402 tccctccacatgctctcagagctcagctcagctaaataatgagcaaaatccagctgagcaga 1461
DB 1587 TCTGCTCTACATGCTCCAGCTTCACTATGACATTAATAAGACAAAATTCATGCGACGGG 1646
QY 1462 aagacaacacagctgctcagatctgctgctgctcagcgagctgcttgaaggacacacgca 1521
DB 1647 AAGAAACATATCAGTGCCCTGTCTCAGTGTGACCGAGGAGCTGTCTGCAAGGCGACACAGG 1706
QY 1522 gtggttaccggttccatcatcatgctgctttagtggtgagcccttggggctgggagacctc 1581
DB 1707 TTGGCTATGAGGTTCCACACTGCTGCTTCCAGTGCATGCCCCCTGTGAAGCTGGGACATTT 1766
QY 1582 ctcaacaagatgacacttaagaatgacagcctgtgtggaagaagagtgagcagctgag 1641

OY	382	gagctcaagaagagacctcttcacattctccctaagtgctgcgacgtatggtgcctgac	441
Db	385	GAGATGCAAGAGAGATCTTCCCAACCACTCCTCCCAAGTGTGGCACTCATGGGCTGAT	444
OY	442	agacacaacgctgtctgcaccacaagcgcgcctctgtagacccttctcgtgccatgatt	501
Db	445	AACCTGACCAACGCTGTCCACCACTGCTGCCCTCTGAGCCCTTTCTATAGCCCTGGTC	504
OY	502	agctatcgccagccagccagcagctcacaagctgtaaacgycagatccctcttccctgyc	561
Db	505	AGCTATGAGCGAGACACACCGTGAATCTCTCATGTGGGAACGCAATTTCCGCTCTTGGCG	564
OY	562	accatccccaatgacaagtaecaggtgtgagaccatgtgtgtcgtctgtcagaaagltcgg	621
Db	565	ACCATCCCCAGCGCATTAAGTACAGAGTGAAGTATAGTGTGCGGCTGTGAGAGCTTGGC	624
OY	622	tggaccctggatctctctgtgttgcagcagtgaaactatagggcagctaggggtcagga	681
Db	625	TGGGTCTGTGATCTCTGCTCGTGTGGCACTATGTGTACAGGGGCACTGGGCTTACAGCG	684
OY	682	cttgagaacacagcgacatgtgtcaagggagctctgattgtcttcaagagataatgaccttc	741
Db	685	CTGAGAGAGCTGGGCACTCCACCGGGCATCTGCTGTGCTTTAAGAGAGTGTGCTTC	744
OY	742	tctgcccaggttggtgcgatatgagaagatgcatgtcctcatatgcacactgtgccaggtgcg	801
Db	745	TCCGCCAGGGGGGTGACCCAAAGATGTGACAGGATGATGTCGTGGCTGAGACAGG	804
OY	802	gcacacgtcgtgtgtgttttttccacgcggcagttgtgccaggtgttttcagagtcgt	861
Db	805	ACCAACCTGTGCTGTGGTCTTCTTAACCGGCACTGGCTGGAGTGTCTTCAAGGCTGTG	864
OY	862	gtgtctgacacaacctgactgtgcgaaggtgtgtgtgcctcaagagccttgagccctccag	921
Db	865	GTCGTGGCCAACTGACTGTGCAAGTGTGATGTGCTTCCGAAGACTGGCCATCTTCACG	924
OY	922	cacatacactgtgggtgtcccggtgacacgcgcatgtggtatgtgtgtcgtgtgcacacag	981
Db	925	TACATCAACCAATGTGCCCGGATTCAGAGGCAATTGGAGCGTGTGGGGTGGCCATCCAG	984
OY	982	aagagggctgtccctgtgcctgtgaagcgcttgaagaagcctatgtcccggtcagacaagag	1041
Db	985	CAGAGCAAGACCTCTGGCCCTGAAGGAGTTGGAAGTCTTAATGCACGACATGTATGGT	1044
OY	1042	gcccctagggccttgcacaaagggctcctcgtgtgacagagaacatcagctctgcagaaatgyc	1101
Db	1045	GCTGCCAAGAACTTGCACAGAGGGGCTCTGTGTGGCACTTACCACTGTGCAAGGAGTGT	1104
OY	1102	caagcttctcatgtracacacgatatgcaccaagctcaaaagccttctcatgattgtctctac	1161
Db	1105	CAGCTTTCAAGCAATGGAAGAACATGTGCCGAGCTTGGAGCTTCTTCATGAGCGCTCCATC	1164
OY	1162	aagcataacgggtctgtgtatgtcgtgtggtcccatatgctctcacacagctcctgtgtctgc	1221
Db	1165	AATGTGATGAGGCTGTGTGTGTGTGTGGCCAAAGGCTCTCACACACTCTCTGGAGTATACC	1224
OY	1222	tcttgagctgttctcacaagggcagagctcaccctctgcagcttttggagagaatccaacag	1281
Db	1225	TCTGTGGACCTGTGCGACAGAGGCCCAAGTCTACCTCTGGCACTTCTTCAGCAATCTACAG	1284
OY	1282	gtgcatctcctcttacaacaaggaacactgtgtgcgtttaaagtaacaagatatccctcagt	1341
Db	1285	GTAATTTCTTCTTACTAATAGAAGACGTATGACATTCGATGACAAAGGGGACCTTACTAGT	1344
OY	1342	agctataacataattgcctctggagctggaatgtgaacccaagtgtgaccttcaaggtcctggt	1401
Db	1345	TATATATGACAACTACCGCTGTGGACTGGAAATGACCTGAATGAGACCTTTGAGGTCTATTGGT	1404
OY	1402	tccctccacatgtgtctcagctcagctcaaaaataaataagaaacaaaatccatggtggacga	1461
Db	1405	TCTGCTCACGTCTTCCAGTTCTCATAGACATTAATTAACAAAATAATCCAGTGGACGGG	1464
OY	1462	aaggacaacacaggtgtccctaagctgtgtgttccagagacgtgtcttgaagggtaccacagga	1521

Db	1465	AAGAA	CATCGAGG	CGCTGTG	GTAGTGA	CTTAC	GAGGAG	CGACTGTCTCGA	AGGACCA	CAGG	1524	
QY	1522	gtgtt	aaaggtttt	caatac	tgtctgt	tttga	tgtgtgtt	gcccc	tgttg	ggcttgg	ggaccttc	1581
Db	1525	TTGGT	CATGGGT	TTC	CCAC	CACTG	CTTG	TGA	GTGATG	CGCTCTGT	GAAGCTTGG	1584
QY	1582	ctcaaca	agagtga	ccctctac	agatgc	caagc	ctcttgt	ggg	aaaga	agagtgg	ggaccttga	1641
Db	1585	CTCAAC	ACAGAGT	GTAGCTT	CCAC	ACTG	CGAC	CTTGT	GGAA	CAAGAA	TAATGGG	1644
QY	1642	ggaag	ccagac	cttctcc	cgag	ctgtgt	ttt	ggctt	gtg	tgag	caac	1701
Db	1645	GGGAG	CTCAG	CGCTCTT	CTC	ACG	CAC	CGTGG	AGTTCTT	GGGGGG	GGGATTA	1704
QY	1702	tgggt	gtcgtcgtg	gaag	taac	caag	cttctgt	cttctgt	cttctgt	ggag	ctgtgtg	1761
Db	1705	TTGGT	GCTATT	ATTAG	ACGCTT	AA	CAGCCTT	ATTC	TCTCTG	CTGTGAT	TGGACTCTG	1764
QY	1762	tttgc	tgtgca	ccctcag	acaccc	ctgtgt	tgtag	tgtag	tcag	gggggg	ggccgc	1821
Db	1765	TTTTG	CTGTGG	CGTTT	TCAC	ACAG	CGCGT	GTGTG	AGGTG	TCAG	GTGGGGT	1824
QY	1822	atgct	ggagctt	ccctctg	cgac	gaagt	gtgtg	cag	ccctctat	tggtctt	tgtgg	1881
Db	1825	ATGCTG	GGGTTC	CTTGTG	TACGTGG	AGATT	GGACG	CTCTAC	AGCTTCTT	CCGG	AAAGCC	1884
QY	1882	agagc	tgtgtgt	ctgtc	gaag	ccag	ggccctt	ttgg	ccctgtgt	tttca	ccactct	1941
Db	1885	GTG	CCCGG	GTCTGT	GCTG	CGCTAC	ACCCCTT	TTTCTG	CTGG	TTTGG	CCATT	1944
QY	1942	tgctg	taagagtc	gtc	taattca	ctaacta	atcat	atcat	cttca	agttt	cttcc	2001
Db	1945	TGTGT	GACAAAT	CGCCTCTT	CCAC	ACTG	GTATCAT	ATCTT	CAAGTTT	CTAC	CAAGT	2004
QY	2002	acatt	ctac	caagc	ctgtgt	gtcc	aaac	caag	tgctgt	gtctt	tgtgt	2061
Db	2005	ACATT	CTAC	CAAC	ACTT	GTGG	CCAA	MAA	CAAG	GTGG	CGGAAT	2064
QY	2062	ggcgg	cccaag	ctgtctat	cttct	taact	tgtgt	gtgt	gtgtg	agccca	ctgtct	2121
Db	2065	ACGGT	CCATT	TTGTTCT	CTGTG	CTC	CAG	CGGCTT	GTCA	TGTGA	CCAC	2124
QY	2122	gaata	ccagc	gtctt	cccc	ccat	ctgt	gtgt	gtgt	gtgtg	tcag	2181
Db	2125	GAGT	TAAG	CGCCTT	CCCC	CCAT	CTGTG	ATTTCTT	GA	TGCAC	AGAGT	2184
QY	2182	ttcata	ctgtg	cttctct	ctctac	aat	tggtctt	ctctt	ctca	atg	tggtctt	2241
Db	2185	TTCTG	GGGTG	CTTTC	GCAC	CA	CAATC	CTCTT	CCAT	TAGAC	CCTTT	2244
QY	2242	cttgg	taag	ggagctt	gtgc	gaag	taac	aa	ggag	ccaa	tggtgt	2301
Db	2245	CTGGT	TAAG	GAAC	TGCG	CGGAG	AACT	ATA	GAAG	CCAA	TGTGT	2304
QY	2302	ttcaac	ttc	gttcc	gtgat	ctgtctt	cttca	cca	ggc	cgctt	cta	2361
Db	2305	CTCC	ACTTTC	GTGAT	CTCTG	GAATC	CGCTT	CTT	CCAT	GTCC	AGATTT	2364
QY	2362	ctgct	gtgg	ggca	caat	gaat	gtgt	gtgt	gtgtg	agag	ctgtg	2421
Db	2365	CTACC	CGGG	TCAT	GTGTG	CGCAG	GGGCTG	CCAC	ACTCTG	ATG	GGCGG	2424
QY	2422	ttctg	ctctaa	gtctat	gtat	ctctt	ctg	cgcc	agag	acttca	acag	2481
Db	2425	TTTC	CTCCT	TAAT	CTCAT	GTGAT	TCTCT	CGCG	TCC	AGAA	CTCA	2484
QY	2482	caggc	ctccat	ctc	agata	ca	aggg	cgct	gtc	ggct	cca	2526
Db	2485	CAGG	CTTC	AT	CTC	AGACT	CTAC	AGAG	CGCTG	CGG	ACT	2529

RESULT 8

AF301161	LOCUS	AF301161	2526 bp	mRNA	linear	ROD 12-FEB-2001
DEFINITION		Mus musculus strain C57BL/6ByJ putative sweet taste receptor T1R1 (Gpr70) mRNA, partial cds.				
ACCESSION		AF301161				
VERSION		AF301161.1	GI:12745517			
KEYWORDS						
SOURCE		house mouse.				
ORGANISM		Mus musculus				
		Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
REFERENCE		1 (bases 1 to 2526)				
AUTHORS		Li,X., Inoue,M., Reed,D.R., Hugue,T., Puchalski,R.B., Tordoff,M.G., Ntironi,Y., Beauchamp,G.K. and Bachmanov,A.A.				
TITLE		High-resolution genetic mapping of the saccharin preference locus (Sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to mouse distal Chromosome 4				
JOURNAL		Mamm. Genome 12 (1), 13-16 (2001)				
MEDLINE		21030739				
REFERENCE		2 (bases 1 to 2526)				
AUTHORS		Li,X., Reed,D.R., Hugue,T., Puchalski,R.B., Tordoff,M.G., Beauchamp,G.K. and Bachmanov,A.A.				
TITLE		Direct Submission				
JOURNAL		Submitted (29-AUG-2000) Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104-3508, USA				
FEATURES		Location/Qualifiers				
source		1..2526				

```

gene
CDS
1..>2526
/gene="Gpr70"
1..>2526
/gene="Gpr70"
/codon_start=1
/product="putative sweet taste receptor T1R1"
/db_xref="GI:12745518"
/translacion="MLFMAHLILSLDLAAYCWAFCSCORTESSPGFSLPGLLAGI
FSLHADIQVHRPLVLCSDRSDSFNGSYLLPAMRTYVEIINNSTALLPNTLGYE
LYDSESSNYSYATLRVLAQGGHLEMQRLDRNLSKVVALLIDPDNDVAITVALIL
SPFLMPLYSANYSATLVILQSGKRPSPFLRTISDKYQVEYVYTRLLQSGRWYISLGS
GDIQOLGQALEELATPRGICVARKNVVPLAQGDPRQCMQMMLLAARTVYVVS
NRHLQGVFRSRVYLANLTKWMAISEMALSTNPVNSGIDGTYLGLVALIQQOVPE
LKEEESYVOAVMGAPRPTCEGSMWCGNIORECHAFETTMPELGAFMSAAAYVEE
AVYAVANGLHOLLIGTSGTCARGPVHQMOLLQITVKNFLHRTKTVAFDQKDGITGY
LIAMDNGNPGTPEFVIGSASLSPVHLIDNKTQIMGKNKQVYVSCTRCLEGGHNR
LYMGSNHCCFECMPCCEAGTFLNLTSELHTCOPGTVEMAPDESSACFSRTVEFLGMHEE
ISVLVLANLTLILLIGTAGLFAMWRLLHTPVRSAGRGLCLMLGSLVAGSCLSEF
GKPVVPACILRQPLFSLGFAFLSCLTIRSFQVLIIFKSTKVPFTFYTMAONAGAGI
FVYSSTYVLELCLTWMAMPRTPREQRPVHILVECTFVNSVGFLVAFANHLIS
IIVPVCSTYIGKELPENYNEAKCVFESLILHVSVALPPTMSIGSLPRAVNVLAGI
ATLSEGSSTYLPKCYVILCRPELANTBHFOASIQDYTRRCGTT"
485 a 755 c 679 g 607 t
BASE COUNT
ORIGIN

```

Query Match	66.48;	Score	1678.4;	DB	10;	Length	2526;
Best Local Similarity	79.38;	Pred. No.	0;				
Matches 1985; Conservative	3;	Mismatches	514;	Indels	0;	Gaps	0;

QY 22 ctgctggccgcgcagcttcacatctctctgctgcgcggcccttgctcgcctacgcagag 81
Db 25 CTGCTCAACCTGCAGCTGGCCGTTGCTTACGCTGGGGCTTTCAGCTCCAAAGACGAA 84
QY 82 tctctcctgcagctcacacctccgcgagatcacctctcgcgcagccctgtctccctccat 141
Db 85 TCCCTCCAGGTTTCAGGCTCCCGGGGACTCTCTCTGGCAGGCCGTGTTCTCCCTCAT 144
QY 142 tctgcctctctgcagctgtgagcacagaccgcgaggtgacccctgtgtgacagctctgagc 201

D	145	GGAGACTGTCGACAGTGTAGACACAGACCTCTGGTGTACAAAGTTGTGACAGTCTGACAG	204
Q	202	ttcaatgagcagagctaccaccctcttccaggcatgcgcttgaggtttgaggagataaac	261
D	205	TTTCAACGGCGCATGGCTATCACCTCTTCCAAACCCATGGCGTTTCAACCGTTGGAGGAATAAC	264
Q	262	aactccaaggccctgcctgcctcccaaatcaacctgggtgacacagctgtaagtgatgtgt	321
D	265	AACTCCACAGCTCTGCTTCCCAACATCACCTGGGGATGATGACTGTAATGACGTGTCTCA	324
Q	322	gactctgcacaatgtgtatgaccagcgtgagagtgctctccctgcaccaggcgacaacacata	381
D	325	GAGTCTTCCAAATGTGTATGTGACACCCCTGAGGGGTGCTCCGCCAGCAAGGAGACCGCACCTA	384
Q	382	gagctccaaggagagacctctccaatatctccctacaggtgtctgcaagtatgttgccctgac	441
D	385	GAGATGCGAGAGAGATCTCCGACACCACTCCCAAGGTGGTGGCACTCATTTGGGCGCTGAT	444
Q	442	agcacaacacgtgtctgcacacacacagccgcccctgtgtgacccttccctgcgtgagccatgat	501
D	445	AACTATTGACACAGCTGTACCACTGCTGCTGCTGTAGCCCTTTTCTGATGCCCCGTGTC	504
Q	502	agctatgtgcgcacagcagcagcagcgtcacgctgaagcggcaglatatccctcttctcgtgc	561
D	505	ACCTATTGAGGGGAGGAGCGGTATCTCTAGTGGGAAGGCAAGTTCCCGCTCTTGGCG	564
Q	562	accatccccaatgataagttaccaggttgagaacacatgtgtctgtctgtcacgaattccgg	621
D	565	ACCATTCGCCAGGATTAATACAGGTGTGAATCTACTGGCGCTGCTGCAAGACTTCGGC	624
Q	622	tggacctggatctctctgtgttggaagaagtgaacatactggacgtcagggtgtcaggaca	681
D	625	TGGGATCTGGAATCTGCTGTGTGGACGCTATGCTATACAGGGCAAGCGTGGGCGATACAGGGC	684
Q	682	ctggagaacacagcgaactggtcagggagatctgcattgtcttccaagacatcatgtccctc	741
D	685	CTGGAGAGAGCTGGCCACTTCACAGGGGATCTGCGTCCGCTTCAAGAAGTGTGCTCTC	744
Q	742	tctgtcccaagtgtgctgatalgagagatgcatgtcctcatctgcgcacactgcccagccggg	801
D	745	TCCGCCCCAGCGGGGAGACCCAAAGGATGACAGCGCATGTGCTGTGGCTGTCTCCAGCAGG	804
Q	802	ggcaccgtctggtgtgtttttccaagccggagttgtgcagggtgtgttttcagtgctccgg	861
D	805	ACCACCGTGGTGTGTGCTTCTTAACCGGACCTGGATGAGTGAAGTCTTCAAGTCTGTG	864
Q	862	gtctgcacacaacactgaactggtgcaaggttgtggtgcgcctcaagaacgtgtgcccctccag	921
D	865	GTCGTGGCCAACTGACTGCGAAGAGTGTGATGCCCTTCCAAAGACTGGGCACTCTCCAGC	924
Q	922	caacatactgtgggtgtgcccgagatcccaagcatcttgagatgtgtgtctgtggcgtgtgccacg	981
D	925	AACATCCCCAAATGTGTCCGGGATCCAGGCGATTGGAGCGGTGGGGGTGGGCATCCAG	984
Q	982	aagagagctgtccctgctgcctgctaaagcggtttgaaagaagcctatgcccggcgagacaaggag	1041
D	985	CAGAGACAAGTCCCTGGGCTTAAGAGATTGTAAGATGCTATGTCCAGGCGAGTGAAGGT	1044
Q	1042	ggcccttagccttgcccaaaagaggtctctgtgtgagcgagcaatcaagctctgcagaagaatgc	1101
D	1045	GCTCCCAAACTCTGCCAGAGGGGTCTCGTGGCGGCACTAACCAAGCTGTGCAAGGAGTGT	1104
Q	1102	caagcttccatgacacacagatgtcccaagctcaaaagccttccatctcatatgattcgtctac	1161
D	1105	CACGCTTTCACAGACTGGAACAAATGCCCGAGTGTGAGACCTTCTCATAAGACGCTGTGCTAC	1164
Q	1162	aacgcataccggcgtgtgtatgtcgtgtggtgccatgtgcctcacacaagctcctgtggctgtgc	1221
D	1165	AATGTATGTAGGCGTGTATGTGTGTGGCCACACGGCTCTCACACAGCTCTCCGGGATTAAC	1224
Q	1222	tctgtagctgtgttccagggtggcgagcttaccctctgtagagtttttgagacagatcccaag	1281
D	1225	TCTGGGACTCTGTGCAGAGGCCCATGCTCACCCCTGGGCGAGTCTTGTACAGAGATCTAACAG	1284


```

Db 2162 AGTCACACTGTAGAGCTTCCCTGTGGCTTACCACACACATCTCCCTCCATCACTA 2221
Oy 2225 ccttgctgaagctactggtgaagacttgccaagaagaactacagaagccaatg 2284
Db 2222 CCTTGCTGAGCTACTGCTGAGTAAGAGCCAGAGACTATATATAGCCAAATGTG 2281
Oy 2285 tcactcaagctctctcactcaactctgctcgtgatcgctctctcaccagcg 2344
Db 2282 TCACCTTAGGCTGCTGCTCAACTTCATCTGATGCGCTTCTTCACATGCGCACCA 2341
Oy 2345 tctcagcagcaagtaactcgtcctcgcgccaacatgagtgtgagctgagcagc 2404
Db 2342 TTTACAGAGGAGCTACTGCTGCGGTGCAATGTGCTGCGAGCGCTGACACACAGAG 2401
Oy 2405 gcgcctcgtggtgtattcttctgctaagtctgaagtctcgtcgcgcagcctca 2464
Db 2402 GCGGCTTAGCGGTACTTCTCCCTCCCAAGTGTATGTATCTCTCCGCTCAGAGCTCA 2461
Oy 2465 acagacagagcaactcagagcctccatcaggaactacagagcgctcgctcagc 2523
Db 2462 ACAATACAGAACATTTCAGGCTTCATCCAGACTACAGAGCGCGCTGCGCACTACG 2520

RESULT 10
AX282896 1101 bp DNA linear PAT 02-NOV-2001
LOCUS AX282896 Sequence 5 from Patent WO0172842.
DEFINITION AX282896
ACCESSION AX282896
VERSION AX282896.1 GI:16609870
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
AUTHORS Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
Abidin,A., Zambrowicz,B. and Sands,A.T.
TITLE Novel human 7tm proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0172842-A 5 04-OCT-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source Location/Qualifiers
1..1101
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 219 a 308 c 323 g 250 t 1 others
ORIGIN
Query Match 39.6%; Score 1001.2; DB 6; Length 1101;
Best Local Similarity 99.6%; Pred. No. 3,1e-207;
Matches 1004; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Db 301 GCGTCAGAGCGCTGGCCCTCTCCAGGACATCACTGGGGTGGCCGGATCCAGGCAAT 360
Oy 955 gggatgtgtctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1014
Db 361 GGGATGTGTCTGGGGTGGGCGCATCCAGAGAGGGCTGTCCCTGGGCTCAAGGCGTTTGA 420
Oy 1015 gaagctatgcccgggagagagagagagagagagagagagagagagagagagagag 1074
Db 421 GAAAGCTTATGCCCCGGGCGAGACAGAGAGCCCTTAGGCTTGGCCAGAGAGGCTCTG 480
Oy 1075 agcagcaatcagctctcagagagagagagagagagagagagagagagagagagagag 1134
Db 481 AGCAGCAATCAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Oy 1135 aaagctctcagatgagatgagatgagatgagatgagatgagatgagatgagat 1194
Db 541 AAAGCTTCTCCATGATGTTGCTGCTACAAAGCATCCGGGCTGTGTATGCGGTGGC 600
Oy 1195 ggcctcagcagctcgtggtggtggtggtggtggtggtggtggtggtggtggtggt 1254
Db 601 GCGCTCCACAGCTCTCTGGGCTGTGCTCTGTGAGCTTGTTCAGAGGCGCGAGTCA 660
Oy 1255 tggcagctcttggagcagatccacaaagtgatctctctcctacacaaagagac 1314
Db 661 TGGCAGCTTTTGGAGCAGATCCAGAGGTCATTTCTTCACAGAGAGAGAGAGAG 720
Oy 1315 tttatgacaaagagagagagagagagagagagagagagagagagagagagag 1374
Db 721 TTTAATGACAAACAGAGATCCCTCAGTAGCTAATTAATTTGCTGGAGCTGGAATGA 780
Oy 1375 cccaagtgagcctcagcagctcgtggtggtggtggtggtggtggtggtggtggt 1434
Db 781 CCCAGTGTGAGCTTACCGGTCTCGGTTCCTCCATATGTCCTCAGTTCAGTAAACATA 840
Oy 1435 aatgagacaaatccagctgagcagagagagagagagagagagagagagagag 1494
Db 841 AATGAGACCAAAATCACTAGTGCGACGAGAAAGACACAGAGTCCATGCTGTGTTCC 900
Oy 1495 agcagctctctgaaagagcagcagcagcagcagcagcagcagcagcagcagc 1554
Db 901 AGCGACTGTCTTGAAGGCGACCGAGAGGTGTACGGGTTTCCATCATCGCTGCTTAG 960
Oy 1555 tgtgtgacctgtgggctggtgagcctctcctcaacaagagtgacctac 1602
Db 961 TGTGTGCTCTGTGGGCTGGAGCTTCTCAACAGAGATTAATCTCTAC 1008

RESULT 11
AL591866 102539 bp DNA linear PRI 25-OCT-2001
LOCUS AL591866 Human DNA sequence from clone Rpl11-58A11 on chromosome 1, complete
DEFINITION AL591866
ACCESSION AL591866
VERSION AL591866.13 GI:16501194
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 102539)
AUTHORS Wallis,J.
JOURNAL Direct Submission
Submitted (25-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequests@sanger.ac.uk
On Oct 26, 2001 this sequence version replaced gi:15022299.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
COMMENT

```


This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C-elegans/wormpep>. This sequence was generated from part of bacterial clone configs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>. Rpl1-58A11 is from the library RPl1-11.1 constructed by the group of Plietier de Jong. For further details see <http://www.chori.org/bacpac/home.htm>. VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RPl1-58A11. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true right end of clone RPl1-58A11 is at 102539 in this sequence. The true left end of clone RPl1-239P22 is at 13247 in this sequence. The true right end of clone RPl1-650H14 is at 2000 in this sequence.

FEATURES

source 1..102539
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RPl1-58A11"
/clone_lib="RPl1-11.1"
582..618
/note="Sequence from overlapping clone RPl1-650H14 (AL158217). Assembly confirmed by restriction digest."
1452..1473
/note="Sequence from overlapping clone RPl1-650H14 (AL158217). Assembly confirmed by restriction digest."
48541..48603
/note="Sequence from overlapping clone RPl1-239P22 (AL159177). Assembly confirmed by restriction digest."
69007..69267
/note="Sequence from overlapping clone RPl1-239P22 (AL159177). Assembly confirmed by restriction digest."
100154..100179
/note="Sequence from overlapping clone RPl1-239P22 (AL159177). Assembly confirmed by restriction digest."
misc_feature 25289 a 26768 c 27360 g 23122 t
BASE COUNT 25289 a 26768 c 27360 g 23122 t
ORIGIN

Query Match 36.9%; Score 932.6; DB 9; Length 102539;
Best Local Similarity 99.9%; Pred. No. 2.6e-192;

Matches 932; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1594 gacctacagatgacagctgtgtggaagaagatggcaccctgaggaagccagacc 1653
|||||
Db 75842 GACCTCTACAGATGCCACTTGTGGAAAGAGTGGCACCTCAGGAACCCAGACC 75901
QY 1654 tgcctccgagcactgtgtgttttgccttgcgtgagacacccctctgtgtgtcgt 1713
|||||
Db 75902 TGTCTCCCGCACCTGTGTGTTTGGCTTTCGTGAGACACCTTGGGCTGCTG 75961
QY 1714 gcaagctaaagctgtgtgtcgtcgtcgtgagacagctgctgtttgcctggac 1773
|||||
Db 75962 GCAAGCTAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 76021
QY 1774 ctgaagaccctgtgtgtgtcagcaggaagcgcctgtgtcttcttctgtcggctcc 1833
|||||
Db 76022 CTAGAACCCTGTGTGTGAGTGTGAGGAGGCGCCCTGTGTCTTATGCTGCGCTCC 76081

QY 1834 ctggcagcaggtagtgagcagcctctatgtctctttgggaaaccacagcctgcgtgc 1893
|||||
Db 76082 CTGGCAGCAGGTAGTGGCAGCCCTGTATGCTTCTTGGGGAACCCAGAGCCTGGTGC 76141
QY 1894 ttgtacgcagagcctcttcttgccttgccttgccttgccttgccttgccttgcct 1953
|||||
Db 76142 TTGCTACGCGCAGGCGCTTGTGGCTTGTGTTCACCATTTCTGCTGCTGACAGTT 76201
QY 1954 cctcaatccaaactaatcaatcattcctcaagtttccacaaagttacattcattccac 2013
|||||
Db 76202 CGCTATTCCTCACTATCATCATCTTCAAGTTTCCACCAAGTACTCATTTCACTACAC 76261
QY 2014 gctcgtgtccaaacacacagctgtgcctgtgttltgtatgataagctcagcggccagctg 2073
|||||
Db 76262 GCTGGGCTCCAAACACACAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 76321
QY 2074 ctatctgtcttaactgt 2133
|||||
Db 76322 CTATCTGTCTTAAGCTTGT 76381
QY 2134 tcccccactgt 2193
|||||
Db 76382 TTCCCATCTGT 76441
QY 2194 ttccctacaatgtgcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2253
|||||
Db 76442 TTCCCTCTCAATGTGCTCCTCTTCCATCATCATGCTTGTGCTGAGCTGCTGCTGAGT 76501
QY 2254 ttgcagagaatctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2313
|||||
Db 76502 TTGCCAGAACTACACAGAGGCGCAATGTGTACCTTGAAGCTGTCTCACTTCGTG 76561
QY 2314 tccctgagctcctctctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 2373
|||||
Db 76562 TCTGTGATGCGCTTCTTTCACACAGGCGCAGCTGTACGAGCGCAATGCTGCGGCC 76621
QY 2374 aacatgtatgt 2433
|||||
Db 76622 AACATATGTGCTGTGCTGTGAGCAGCCCTGACAGGCGCTGTGTGTGTGTGTGTGT 76681
QY 2434 tgcctagctatcctctgtgc 2493
|||||
Db 76682 TGTCTACGTATCTCTTGTGCGGCCGACAGCTCTACAGCAGACGACATCTCCAGGCTTCAT 76741
QY 2494 caaggtacacagagcgcctgcgtgcctcactga 2526
|||||
Db 76742 CAGGACTACACAGGCGGCTGCGGCTCCACCTGA 76774
RESULT 12
AC108123 191481 bp DNA linear HTG 25-JAN-2002
LOCUS AC108123/C
DEFINITION Homo sapiens chromosome 16 clone RPl1-56511, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
ACCESSION AC108123.1 GI:18369969
VERSION AC108123.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 191481)
DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 191481)
DOE Joint Genome Institute.
AUTHORS Direct Submission
TITLE Submitted (25-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI

for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 186 a 260 c 294 g 198 t

Query Match 14.6%; Score 369; DB 10; Length 938;
Best Local Similarity 99.8%; Pred. No. 3.4e-171;
Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 367 gagcaaccacacatagactccaagaagaccttccactatccctcagtgctgca 426
DB 141 GGGCAACACACATAGAGCTCCAGAGACCTTCTCCACTATCCCTACGGCTGCA 200
OY 427 gtgattggcctgagacacacacgctgtccacacacgcccctgtgagcccttc 486
DB 201 GTGATTGGCCTGACGACCAACCGTCTGCCACACACGCCCTGCTGAGCCCTTTC 260
OY 487 ctggtgcccattgattagctatgctgagcagcagcagcagcagcagcagcagat 546
DB 261 CTGGTGCCCATGATGATGATGAGGGCCAGAGAGAGCCTCAGCTGAAGCGCAGTAT 320
OY 547 ccccttttctgagacacatcccccaatgacaaagtacaggtgagacatgtgtctg 606
DB 321 CCTCTTCTTCTCGCACACCTCCCAATGACCAAGTACAGGTGAGACCAATGCTGCTG 380
OY 607 ctgcagaattcgggtgagcctgactctctgtgtgagcagcagcagcagcagcag 666
DB 381 CTGCAGAAATTGGGGGAGCCTGATCTCTGTGTTGGCAGCAGTACACATATGGGCA 440
OY 667 ctagggtgagcagcactgtagaagacacagcagcagcagcagcagcagcagcag 726
DB 441 CTAGGGGTGACAGGCACTGAGAACACAGCGACAGTGTGACGCTGCAATGCTTCAAG 500
OY 727 gacatcatgccccttctctgcccagtggtgagtgatgagaagatgacagtcctatg 786
DB 501 GACATCATGCCCTTCTCTCTCCAGGTGGGAGTGAAGAGATGATGATGCTCATGCGCAC 560

RESULT 2
A1742401/c 562 bp mRNA linear EST 19-DEC-1999
LOCUS A1742401
DEFINITION wg40e02.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:2367578 3' similar to TR:093557 093557 PUTATIVE ODORANT
RECEPTOR ;, mRNA sequence.
ACCESSION A1742401
VERSION A1742401.1 GI:5110689
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 562)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 840 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 468.
Location/Qualifiers
1..562
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2367578"
/clone_id="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"

FEATURES
Source

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and six circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NBHSF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NB2HB-9W pool 1: 758280-760583, 772104-774407 Soares NBHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 116 a 161 c 167 g 117 t 1 others

Query Match 11.8%; Score 299; DB 9; Length 562;
Best Local Similarity 99.7%; Pred. No. 1.3e-136;
Matches 349; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2177 ttggtcttaactatgctctctctacatgagcctctccatagtccttgcctgca 2236
DB 531 TGGGCTTCAATAGCGGCTTCTCTTACAAATGGCTCTCTTCCATCAATGCTTGGCTTGA 472
OY 2237 gctactctggttaagagcttgcagagaactacaaagagcgaatggtacactttagc 2296
DB 471 GCTACTGGGTAAAGACTTGGCCAGAACTACACAGAGCCAAATGTGCACCTTAAGC 412
OY 2297 tgccttcaactcgtgtcctgagtcgcttcttccacacagcagcagcagcagcag 2356
DB 411 TGCTTTCAACTTCGTCCTGATGCGCTTCTTACACAGGCGACGCTACAGAGGCA 352
OY 2357 agtactcgtcctgagcacaatgagtgctggtgagcagcctgagcagcagccttg 2416
DB 351 AGTACCTGCTCGGCGCAACATGATGGCTGTGAGCAGCCGTGACACCGGCTTGGTG 292
OY 2417 ggtatttctgacctagtgctacgtatcctctgcccagcagcctcacaagacagagc 2476
DB 291 GGTATTTCCTGCTTAAGTGTCTACGTATCTCTGCGGCCAGACCTCAACAGCAGAGC 232
OY 2477 acttccagcctccatcagactacagcagcagcagcagcagcagcagcagcagc 2526
DB 231 ACTTCCAGGCTTCCATTACAGACTACACAGAGGGCTGCGGCTCCACCTGA 182

RESULT 3
AA853967/c 496 bp mRNA linear EST 31-DEC-1998
LOCUS AA853967
DEFINITION aj51e10.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:1393866
3' similar to SW:CA58_RAT P48442 EXTRACELLULAR CALCIUM-SENSING
RECEPTOR PRECURSOR ;, mRNA sequence.
ACCESSION AA853967
VERSION AA853967.1 GI:2941505
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 496)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

FEATURES
Source

BASE COUNT	122 a	153 c	134 g	104 t	1 others
ORIGIN					

Query Match 1.7%; Score 43; DB 12; Length 514;
 Best Local Similarity 100.0%; Pred. No. 6.9e-10;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1543 tgcgtcttgagtggtgcccctggagcctctca 1585
 |||||||
 Db 280 tgcgtcttgagtggtgcccctggagcctctca 238

RESULT 6
 A0631494/c 193 bp DNA linear GSS 17-JUN-1999
 LOCUS RPCI-11-473E18.TV RPCI-11 Homo sapiens genomic clone RPCI-11-473E18
 DEFINITION , DNA sequence.
 ACCESSION A0631494
 VERSION A0631494.1 GI:5094129
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 193)
 Zhao, S., Adams, M.D., Niernan, W., Malek, J., de Jong, P. and Venter
 , J.C.
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 Map Building
 Unpublished (1997)
 Contact: Shaying Zhao, William Niernan, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@edj.med.buitalo.edu). Clones may be purchased from
 BACRAC Resources (<http://bacpac.med.buitalo.edu/ordering>) or from
 Research Genet cs (<http://inforesgen.com>). BAC end search page:
http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: 17
 Class: BAC ends.
 FEATURES
 Location/Qualifiers
 1. 193
 /organism="Homo sapiens"
 /db_xref="GDB:7681361"
 /db_xref="taxon:9606"
 /clone="RPCI-11-473E18"
 /clone_1lb="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPCI11 Human Male BAC Library"

BASE COUNT 36 a 66 c 52 g 39 t
 ORIGIN

Query Match 1.1%; Score 27; DB 12; Length 193;
 Best Local Similarity 100.0%; Pred. No. 0.046;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1543 tgcgtcttgagtggtgcccctggag 1569
 |||||||
 Db 193 tgcgtcttgagtggtgcccctggag 167

RESULT 7
 B1580014/c 473 bp mRNA linear EST 06-SEP-2001
 LOCUS B1580014
 DEFINITION R74133.5prtime RE Drosophila melanogaster normalized Embryo pRc-1
 Drosophila melanogaster cDNA clone R74133 5 similar to CG18543:

FBan0018543 GO:[] located on: 3L 66C12-66C12.: 05/21/2001, mRNA
 sequence.
 B1580014
 VERSION B1580014.1 GI:15471436
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 473)
 Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
 , J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
 , R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Mistra, S.,
 Mungall, C.J., Nuno, J., Pacle, J., Parag, V., Park, S.,
 Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin
 , G.M.
 BDGP/HMI RE Drosophila EST Project
 Unpublished (2001)
 Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_estfruitfly.berkeley.edu
 hit genomic AE003555: arm:3L [8254613,8536879]
 estimated-cyto:66C8-66D9: 05/21/2001
 Plate: RE 741 row: C column: 9
 High quality sequence stop: 449.
 FEATURES
 Location/Qualifiers
 1. 473
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="R74133"
 /clone_1lb="RE Drosophila melanogaster normalized Embryo
 pRc-1"
 /sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pRc1; Site_1: XhoI; Site_2:
 BamHI; Library was kindly generated by Piero Carinai at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT 135 a 140 c 111 g 87 t
 ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 473;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1725 gctgcgtcgtcgtcgtcgtgga 1750
 |||||||
 Db 328 gctgcgtcgtcgtcgtcgtgga 303

RESULT 8
 B1485890 522 bp mRNA linear EST 28-AUG-2001
 LOCUS B1485890/c
 DEFINITION RB69243.5prtime RE Drosophila melanogaster normalized Embryo pRc-1
 Drosophila melanogaster cDNA clone RB69243 5 similar to CG18543:
 FBan0018543 GO:[] located on: 3L 66C12-66C12.: 05/16/2001, mRNA
 sequence.
 B1485890
 VERSION B1485890.1 GI:15325673
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 522)
 Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson

J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S., Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin, G.M.

TITLE BDGP/HHMI RE Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT Contact: Stapleton, M.

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
hit genomic: AE003555: arm:3L [8254613,8536879]
estimated-cyto: 66C8-66D9: 05/16/2001
Plate: RE.189 row: D column: 7
High quality sequence stop: 509.

FEATURES
source location/Qualifiers
1..522
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE69243"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFLC-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/note="Organ: embryo; Vector: pFLC1, Site_1: XhoI, Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."

BASE COUNT 146 a 155 c 123 g 98 t
ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 522;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgctgga 1750
|||||
Db 328 GCTGCTGCTGCTGCTGCTGCTGGA 303

RESULT 9
LOCUS BH093096/c 522 bp DNA linear GSS 18-JUL-2001
DEFINITION RPCI-24-357D16.TV RPCI-24 Mus musculus genomic clone RPCI-24-357D16
, DNA sequence.
ACCESSION BH093096
VERSION BH093096.1 GI:14913001
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 522)
Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akincet, B., Levins, M.,
Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gedregeorgis, E.,
Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSSs: RPCI-24-357D16.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 357 row: D column: 16
Seq primer: 77
Class: BAC ends.

FEATURES
source location/Qualifiers
1..522
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-357D16"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTRABAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTRABAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT 152 a 137 c 120 g 113 t
ORIGIN

Query Match 1.0%; Score 26; DB 12; Length 522;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgctgga 1750
|||||
Db 141 GCTGCTGCTGCTGCTGCTGCTGGA 116

RESULT 10
LOCUS B1213517/c 536 bp mRNA linear EST 11-JUL-2001
DEFINITION RE18934.Sprline RE Drosophila melanogaster normalized Embryo pFLC-1
Drosophila melanogaster cDNA clone RE18934 5 similar to CG18543:
F8an0018543 'cell cycle regulator' located on: 3L 66C12-66C12.1;
04/12/2001, mRNA sequence.
ACCESSION B1213517
VERSION B1213517.1 GI:14691241
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 536)
Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
J., Champe, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S.,
Mungall, C.J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Celniker, S. and Rubin
G.M.
BDGP/HHMI RE Drosophila EST Project
Unpublished (2001)
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
hit genomic: AE003555: arm:3L [8254613,8536879]
estimated-cyto: 66C8-66D9: 04/12/2001
Plate: RE.189 row: C column: 10
High quality sequence stop: 524.
location/Qualifiers
1..536
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE18934"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFLC-1"

/sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pFic1; Site:1: XhoI; Site:2:
 BamHI; Library was kindly generated by Piero Carinici at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."
 BASE COUNT 146 a 158 c 128 g 104 t
 ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 536;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1725 gctgcctcctcctcctcctgga 1750
 ||||||||||||||||||||||||||||
 Db 329 GCTGCTGCTGCTGCTGCTGGA 304

RESULT 11
 B1235687 536 bp mRNA linear EST 12-JUL-2001
 LOCUS B1235687/c
 DEFINITION RE31596.5prime RE Drosophila melanogaster normalized Embryo pFic-1
 Drosophila melanogaster cDNA clone RE31596 5 similar to CG18543:
 FBan0018543 'cell cycle regulator' located on: 3L 66C12-66C12;;
 05/12/2001, mRNA sequence.
 ACCESSION B1235687
 VERSION B1235687.1 GI:14704130
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 536)
 Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B., Carlson
 J., Champagne, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George
 R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Misra, S.,
 Munhall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
 Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and Rubin
 G. M.
 BDFG/HMMI RE Drosophila EST project
 JOURNAL Unpublished (2001)
 COMMENT BDFG
 Contact: Stapleton, M.
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
 hit genomic A800355: arm:3L [8254613,8536879]
 estimated-cyto:66C8-66D9: 05/12/2001
 Plate: RE.315 row: H column: 12
 High quality sequence stop: 494.

FEATURES
 source
 1..536
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="RE31596"
 /clone_1ib="RE Drosophila melanogaster normalized Embryo
 pFic-1"
 /sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pFic1; Site:1: XhoI; Site:2:
 BamHI; Library was kindly generated by Piero Carinici at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT 148 a 157 c 128 g 103 t
 ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 536;

Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1725 gctgcctcctcctcctcctgga 1750
 ||||||||||||||||||||||||||||
 Db 329 GCTGCTGCTGCTGCTGCTGGA 304

RESULT 12
 A2816179 543 bp DNA linear GSS 20-FEB-2001
 LOCUS A2816179/c
 DEFINITION 2M0084L13R Mouse 10kb plasmid UGCGM library Mus musculus genomic
 clone UGCG2M0084L13 R, DNA sequence.
 ACCESSION A2816179
 VERSION A2816179.1 GI:12986087
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 543)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 1000 Std Error: 0.00
 Plate: 0084 row: L column: 13
 Seq primer: CACACGAGAAACAGCATGACCC
 Class: plasmid ends
 High quality sequence stop: 543.

FEATURES
 source
 1..543
 Location/Qualifiers

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UGCG2M0084L13"
 /clone_1ib="Mouse 10kb plasmid UGCGM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42ny; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (g11473211419b/AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 BASE COUNT 152 a 139 c 105 g 147 t
 ORIGIN

Query Match 1.0%; Score 26; DB 12; Length 543;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgctgga 1750
 ||||||||||||||||||||||||||||
 Db 197 gctgctgctgctgctgctgctgctgga 172

RESULT 13
 B1577858/c 585 bp mRNA linear EST 06-SEP-2001
 LOCUS RE1616.5prime RE Drosophila melanogaster normalized Embryo pf1c-1
 DEFINITION Drosophila melanogaster cDNA clone RE1616 5 similar to CG18543:
 Fban0018543 GO:[] located on: 3L 66C12-66C12; 05/17/2001, mRNA
 sequence.
 ACCESSION B1577858
 VERSION B1577858.1 GI:15469280
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 585)
 AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
 J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
 R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
 Mungall,C.J., Nunoo,J., Pacled,J., Paragas,V., Park,S.,
 Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
 G.M.
 TITLE BDGP/HMI RE Drosophila EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: Stapleton, M.
 BDCP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 hit genomic AE003555: arm:3L [8254613,8536879]
 estimated-cyto:66C8-66D9: 05/17/2001
 Plate: RE:716 row: B column: 4
 High quality sequence stop: 523.
 Location/Qualifiers
 1..585
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_id="RE1616"
 /clone_1lb="RE Drosophila melanogaster normalized Embryo
 pf1c-1"
 /sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pf1c1; Site_1: XhoI; Site_2:
 BamHI; Library was kindly generated by Piero Carninci at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT 162 a 169 c 141 g 113 t

ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 585;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgctgga 1750
 ||||||||||||||||||||||||||||
 Db 328 gctgctgctgctgctgctgctgctgga 303

RESULT 14
 B1366613/c 586 bp mRNA linear EST 01-AUG-2001
 LOCUS B1366613

DEFINITION RE52188.5prime RE Drosophila melanogaster normalized Embryo pf1c-1
 Drosophila melanogaster cDNA clone RE52188 5 similar to CG18543:
 Fban0018543 'cell cycle regulator' located on: 3L 66C12-66C12;:
 05/14/2001, mRNA sequence.
 ACCESSION B1366613
 VERSION B1366613.1 GI:15062641
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 586)
 AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
 J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
 R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mista,S.,
 Mungall,C.J., Nunoo,J., Pacled,J., Paragas,V., Park,S.,
 Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
 G.M.
 TITLE BDGP/HMI RE Drosophila EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: Stapleton, M.
 BDCP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 hit genomic AE003555: arm:3L [8254613,8536879]
 estimated-cyto:66C8-66D9: 05/14/2001
 Plate: RE:521 row: H column: 4
 High quality sequence stop: 524.
 Location/Qualifiers
 1..586
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_id="RE52188"
 /clone_1lb="RE Drosophila melanogaster normalized Embryo
 pf1c-1"
 /sex="male and female"
 /dev_stage="0-24 hours mixed stage embryonic"
 /lab_host="DH5-alpha Tona"
 /note="Organ: embryo; Vector: pf1c1; Site_1: XhoI; Site_2:
 BamHI; Library was kindly generated by Piero Carninci at
 the RIKEN. The library was normalized and excised using
 Cre recombinase. Plasmid cDNA library."

BASE COUNT 162 a 169 c 142 g 113 t

ORIGIN

Query Match 1.0%; Score 26; DB 10; Length 586;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1725 gctgctgctgctgctgctgctgctgga 1750
 ||||||||||||||||||||||||||||
 Db 328 gctgctgctgctgctgctgctgctgga 303

RESULT 15
 BF500786/c 611 bp mRNA linear EST 19-APR-2001
 LOCUS BF500786
 DEFINITION AT15818.5prime AT Drosophila melanogaster adult testes pOTB7
 Drosophila melanogaster cDNA clone AT15818 5 similar to CG13798:
 Fban0013798 located on: 3L 62E4-62E4; 04/09/2001, mRNA sequence.
 ACCESSION BF500786
 VERSION BF500786.2 GI:13692625
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 611)

AUTHORS

Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman, B., Carlson, J., Champagne, M., Chavez, C., Chew, M., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phoumenavong, S., Wan, K., Yu, C., Lewis, S. E., Celisner, S., and Rubin, G. M.

TITLE

BDGP/HHMI AT Drosophila EST Project

JOURNAL

Unpublished (2000)

COMMENT

On Dec 6, 2000 this sequence version replaced gl:11584087.
Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03474: arm:3r [2120287,2430124]

estimated-cyto:62c4-62E5: 04/09/2001

Place: AT.158 row: B column: 6

High quality sequence stop: 571.

FEATURES

source

1. 611
Location/Qualifiers

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="AT15818"

/clone_lib="AT Drosophila melanogaster adult testes potB7"

/sex="male"

/dev_stage="0-3 day old Ore-R males"

/lab_host="Plates AT.10-AT.120: DHS-alpha. Plates

AT.121-AT.319: DHS-alpha Toma"

/note="Organ: ADULT testes; Vector: potB7; Site_1: EcORI;

Site_2: XhoI; The mRNA for the testis library was made

from testes and seminal vesicles hand dissected from 0-3

day old Ore-R males. RNA kindly provided by the lab of

Margaret Fuller. Sized fractionated cDNAs were directly

ligated into potB7. Plasmid cDNA library."

169 a 190 c 164 g 88 t

BASE COUNT
ORIGIN

Query Match

1.0%; Score 26; DB 10; Length 611;

Best Local Similarity 100.0%; Pred. No. 0.19;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1725 gctgctgctgctgctgctgctgga 1750

|||||

DB 345 GCTGCTGCTGCTGCTGCTGCTGGA 320

Search completed: June 11, 2002, 08:20:22
Job time: 5524 sec

OTHER INFORMATION:

OTHER INFORMATION: repeat sequence"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 104..252
OTHER INFORMATION: /note="Nucleotides 104-252 are
OTHER INFORMATION: unique flanking sequence"
US-08-623-906A-1

Query Match 0.9%; Score 23; DB 2; Length 252;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1724 ccgtgctgctgctgctgctgct 1746
|||||
Db 27 ccgtgctgctgctgctgctgct 49

RESULT 2
US-08-863-639A-31/c
; Sequence 31, Application US/08863639A
; Patent No. 5981185
; GENERAL INFORMATION:
; APPLICANT: Matson, Robert S.
; APPLICANT: Coassin, Peter J.
; APPLICANT: Rampal, Jang B.
; APPLICANT: Caskey, C. T.
; TITLE OF INVENTION: OLIGONUCLEOTIDE REPEAT ARRAYS
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, 9th Floor
; CITY: Pasadena
; STATE: CA
; COUNTRY: USA
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Corel Wordperfect 8 version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/863,639A
; FILING DATE: May 28, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Joseph E. Muehl
; REGISTRATION NUMBER: 20,532
; REFERENCE/DOCKET NUMBER: 11859-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 796-4000
; TELEFAX: (626) 795-6321
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid
US-08-863-639A-31

Query Match 0.9%; Score 22; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1724 ccgtgctgctgctgctgctgct 1745
|||||
Db 27 ccgtgctgctgctgctgctgct 6

RESULT 3
US-09-253-691-3/c

; Sequence 3, Application US/09253691
; Patent No. 6124100
; GENERAL INFORMATION:
; APPLICANT: Dong Kyu JIN
; TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
; FILE REFERENCE: 1942/36
; CURRENT APPLICATION NUMBER: US/09/253,691
; CURRENT FILING DATE: 1999-02-22
; EARLIER APPLICATION NUMBER: KR 98-6,278
; EARLIER FILING DATE: 1996-02-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Wordperfect 6.1/Windows
; SEQ ID NO 3
; LENGTH: 397
; TYPE: DNA
; ORGANISM: human
US-09-253-691-3

Query Match 0.9%; Score 22; DB 3; Length 397;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1724 ccgtgctgctgctgctgctgct 1745
|||||
Db 357 ccgtgctgctgctgctgctgct 336

RESULT 4
US-08-831-399-3
; Sequence 3, Application US/08831399
; Patent No. 6312916
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core
; NUMBER OF SEQUENCES: 16
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,399
; FILING DATE: 1-April-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6312916man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 44..499
OTHER INFORMATION:
OTHER INFORMATION: positions 44..115 correspond to sig
US-08-831-399-3 peptide and 116..499 to mat peptide

Query Match 0.9%; Score 22; DB 4; Length 604;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1724 cgctgctgctgctgctgct 1745
|||||
DB 63 ccctgctgctgctgctgct 84

RESULT 5
US-09-228-986-38
Sequence 38, Application US/09228986
Patent No. 6359198
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Niels
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and their use in the Modification of Plant Cell Signalling
FILE REFERENCE: 11000/1020
CURRENT APPLICATION NUMBER: US/09/228.986
CURRENT FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 130
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 38
LENGTH: 822
TYPE: DNA
ORGANISM: Eucalyptus grandis
US-09-228-986-38

Query Match 0.9%; Score 22; DB 4; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1724 cgctgctgctgctgctgct 1745
|||||
DB 132 cgctgctgctgctgctgct 153

RESULT 6
US-08-240-124-1
Sequence 1, Application US/08240124
Patent No. 5516658
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240.124
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: hek-L A2
FEATURE:
NAME/KEY: CDS
LOCATION: 83..799
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 83..139
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 140..796
US-08-240-124-1.

Query Match 0.9%; Score 22; DB 1; Length 1037;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1724 cgctgctgctgctgctgct 1745
|||||
DB 96 ccctgctgctgctgctgct 117

RESULT 7
US-08-453-943-1
Sequence 1, Application US/08453943
Patent No. 5738844
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
APPLICANT: CERRETTI, DOUGLAS P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/453,943
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/240,124
FILING DATE: 03-MAY-1994
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: hek-L A2
FEATURE:
NAME/KEY: CDS
LOCATION: 83..799
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 83..139
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 140..796
US-08-453-943-1

Query Match 0.9%; Score 22; DB 1; Length 1037;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgctgctgctgctgctgct 1745
|||||
Db 96 CGCTGCTGCTGCTGCTGCT 117

RESULT 8
US-09-057-121-1
Sequence 1, Application US/09057121
Patent No. 5969110
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,121
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/240,124
FILING DATE:
APPLICATION NUMBER: US 08/161,132
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/114,426
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/109,745
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SEESE, KATHRYN A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2814-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: hek-L A2
FEATURE:
NAME/KEY: CDS
LOCATION: 83..799
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 83..139
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 140..796
US-09-057-121-1

Query Match 0.9%; Score 22; DB 2; Length 1037;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgctgctgctgctgctgct 1745
|||||
Db 96 CGCTGCTGCTGCTGCTGCT 117

RESULT 9
US-09-358-734-1
Sequence 1, Application US/09358734
Patent No. 6274117
GENERAL INFORMATION:
APPLICANT: BECKMANN, M. P.
TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
TITLE OF INVENTION: RECEPTOR HEK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON

TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1776 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1115
; US-08-531-927B-1

Query Match 0.9%; Score 22; DB 2; Length 1776;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgctgctgctgctgctgct 1745
Db 988 CGCTGCTGCTGCTGCTGCT 967

RESULT 12
US-09-041-886-12/C
; Sequence 12, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Shiroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; NUMBER OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1776 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1116
; US-09-041-886-12

Query Match 0.9%; Score 22; DB 4; Length 1776;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1724 cgctgctgctgctgctgct 1745

Db 988 CGCTGCTGCTGCTGCTGCT 967

RESULT 13
US-08-753-007A-5/C
; Sequence 5, Application US/08753007A
; Patent No. 6074841
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/753,007A
; FILING DATE: 19-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 664...1883
; OTHER INFORMATION:
; US-08-753-007A-5

Query Match 0.9%; Score 22; DB 3; Length 1884;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgctgctgctgctgctgct 1745
Db 65 CGCTGCTGCTGCTGCTGCT 44

RESULT 14
US-09-398-496-5/C
; Sequence 5, Application US/09398496
; Patent No. 6133423
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
; AND USES THEREFOR

NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/398,496
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/753,007
FILING DATE: 19-NOV-1996
APPLICATION NUMBER: 08/699,591
FILING DATE: 19-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07334/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1884 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 664..1883
OTHER INFORMATION:
US-09-398-496-5

Query Match 0.9%; Score 22; DB 3; Length 1884;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1724 cgcctgctgctgctgctgct 1745
|||||
DB 65 CGCTGCTGCTGCTGCTGCT 44

RESULT 15
US-08-261-822A-7/c
Sequence 7, Application US/08261822A
Patent No. 5650553
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph R. et al.
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553tris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261,822A
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2234 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-261-822A-7

Query Match 0.9%; Score 22; DB 1; Length 2234;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1725 gctgctgctgctgctgctgt 1746
|||||
DB 1970 GCTGCTGCTGCTGCTGCTT 1949

Search completed: June 11, 2002, 09:14:02
Job time: 8194 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

```
Run on:      June 11, 2002, 06:49:22 ; Search time 3107.22 Seconds
              (without alignments)
              17012.130 Million cell updates/sec
```

Title: US-09-819-946-1
Perfect score: 2526
Sequence: 1 atgctgtctgtcaagctcg.....ggcgtctgcgtctcaactga 2526

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

word size :

Total number of hits satisfying chosen parameters: 3595312

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Listing first 45 summaries

Database :

1:	gb_emb1.*
2:	gb_htg.*
3:	gb_in.*
4:	gb_on.*
5:	gb_ov.*
6:	gb_pat.*
7:	gb_ph.*
8:	gb_pl.*
9:	gb_pr.*
10:	gb_ro.*
11:	gb_sts.*
12:	gb_sy.*
13:	gb_un.*
14:	gb_vl.*
15:	em_da.*
16:	em_fun.*
17:	em_hum.*
18:	em_in.*
19:	em_mu.*
20:	em_om.*
21:	em_or.*
22:	em_ov.*
23:	em_pat.*
24:	em_ph.*
25:	em_pl.*
26:	em_ro.*
27:	em_sts.*
28:	em_un.*
29:	em_vl.*
30:	em_htg_hum.*
31:	em_htg_iny.*
32:	em_htg_inv.*
33:	em_htgq_iny.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description

C	1	2526	100.0	22561	6	AX282892	AX282892	Sequence
C	2	2526	100.0	22951	6	AX282900	AX282900	Sequence
C	3	2332	91.9	22759	6	AX354027	AX354027	Sequence
C	4	2252	90.7	22252	6	AX282894	AX282894	Sequence
C	5	9599	39.5	1101	6	AX282896	AX282896	Sequence
C	6	882	34.9	102539	2	AL591866	Human DNA	
C	7	882	34.9	191481	2	AC108123	Human sapi	
C	8	705	27.9	705	6	AX282898	AX282898	Sequence
C	9	549	21.7	185061	2	AC062024	Human sapi	
C	10	270	10.7	2270	11	G09859	human SRS C	
C	11	160	7.1	3106	9	HSHER3501	U55324 Human Krupp	
C	12	172	6.8	185056	2	AC062024	AC062024 Homo sapi	
C	13	68	2.7	191481	2	AC108123	AC108123 Homo sapi	
C	14	32	1.3	2520	10	AF127389	AF127389 Rattus no	
C	15	31	1.2	5491	6	AX251144	AX251144 Sequence	
C	16	27	1.1	2526	6	AF301161	AF301161 Mus muscu	
C	17	27	1.1	2526	10	AF301162	AF301162 Mus muscu	
C	18	27	1.1	2526	10	AY032622	AY032622 Mus muscu	
C	19	27	1.1	2892	10	AF337040	AF337040 Mus muscu	
C	20	27	1.1	198306	2	AL611927	AL611927 Mus muscu	
C	21	26	1.0	224	8	AY022959	AY022959 Oryza sat	
C	22	26	1.0	656	3	DM003288	U03288 Drosophila	
C	23	26	1.0	1065	3	AY061524	AY061524 Drosophila	
C	24	26	1.0	17864	2	AC014995	AC014995 Drosophila	
C	25	26	1.0	30894	2	AC108346	AC108346 Rattus no	
C	26	26	1.0	50885	2	AC014946	AC014946 Drosophila	
C	27	26	1.0	62490	2	AC110265	AC110265 Mus muscu	
C	28	26	1.0	127362	2	AC010014	AC010014 Drosophila	
C	29	26	1.0	158805	2	AC104847	AC104847 Oryza sat	
C	30	26	1.0	167909	2	AC098184	AC098184 Rattus no	
C	31	26	1.0	169730	3	AC010006	AC010006 Drosophila	
C	32	26	1.0	171697	2	AC093548	AC093548 Drosophila	
C	33	26	1.0	174706	3	AC010041	AC010041 Drosophila	
C	34	26	1.0	282267	3	AE003555	AE003555 Drosophila	
C	35	26	1.0	309838	3	AE003474	AE003474 Drosophila	
C	36	25	1.0	770	11	AF021116	AF021116 Homo sapi	
C	37	25	1.0	906	6	ZMA01025	AF010295 Zea mays	
C	38	25	1.0	913	8	ZMGST3	X06755 Maize mRNA	
C	39	25	1.0	944	8	ZMA01026	AF010296 Zea mays	
C	40	25	1.0	4668	9	HSY18265	Y18265 Homo sapien	
C	41	25	1.0	5914	9	HSZPHSAL1	X58833 Homo sapien	
C	42	25	1.0	25975	2	AC015255	AC015255 Drosophila	
C	43	25	1.0	69986	2	AC020283	AC020283 Drosophila	
C	44	25	1.0	89285	2	DMB33H4	AL121853 Drosophila	
C	45	25	1.0	167089	3	AC023747	AC023747 Drosophila	

ALIGNMENTS

RESULT	1								
LOCUS	AX282892		2526 bp	DNA	Linear	PAT 02-NOV-2001			
DEFINITION	Sequence 1 from Patent WO0172842.								
ACCESSION	AX282892								
VERSION	AX282892.1		GI:16609868						
KEYWORDS									
SOURCE	human.								
ORGANISM	Homo sapiens								
	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
REFERENCE	1 (sites)								
AUTHORS	Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C., Abidin,A., Zambrowicz,B. and Sands,A.T.								
TITLE	Novel human 7tm proteins and polynucleotides encoding the same								
JOURNAL	Patent: WO 0172842-A 1 04-Oct-2001;								
FEATURES	Lexicon Genetics Incorporated (US)								
source	location/Qualifiers								
	1..2526								
	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
BASE COUNT	469 a	770 c	697 g	587 t	3 others				
ORIGIN									

OY	1021	taagccggggaagaacaagaggcccccaggcccttgccaagaaggccctcggtgagcaagc	1080
Db	1021	TATGCGGGGAGCAAGAGAGAGGCCCTTAGGCGCTTGCCACAAAGGGCTCGTGGACAGAC	1080
OY	1081	aatcagcccttcagaagaatgcgaagccttcattatgircacacagatgccaagctcacaagcc	1140
Db	1081	AATGAGCTCTCAGAGAATGACAAAGCTTTCATGRCACACACGATGGCCAAAGCTCAAAGCC	1140
OY	1141	ttctccatgagttctgccttaacaagcatcacgggctgtgtagcggtggtgccatgagccctc	1200
Db	1141	TTCTCCATGAGATTCTGCGCTACCAAGCATACCGGGCGTGTATGCGGTGGCCATGGCCCTC	1200
OY	1201	cacagagccctcgaggctgtgctccttgagagctgtgtccaaagggccagagctcaaccctggag	1260
Db	1201	CACCAAGCTCTCGGGCTGTGCTCTGGAGCTGTGTTCCAGGGGCCGAGTCTACCCCTGGCAG	1260
OY	1261	cttttgagcagatccacaagaagtgcatttcccttctacaacaagaagacatgtgtgctttaat	1320
Db	1261	CTTTTGGAGCGAGATCCACAAGAGTGCAATTTCTTCACAAAGACACTGTGGCGTTTAAAT	1320
OY	1321	gacacagagatccccccatgactataataatatgtccttgaggactggaaatgagaccag	1380
Db	1321	GACCAACAGATCCCCCTCAGTAGCTATTAACATTAATTTGCTGGAGCTGAGATGACCCAG	1380
OY	1381	tgaagcttcaagcgctcctcggtctccccaatgtctccagttccagctaaacataatatg	1440
Db	1381	TGGAAGCTTCAAGGTCGTGGTCTCTCCACATGCTGCCAGTTCCAGGTAAACATAATATAG	1440
OY	1441	accacaaatccagtgtagcagcggaagaagacacacagtgacctaaagtctgtgtgtccagcgac	1500
Db	1441	ACCAAAATCCAGTGGCAAGGAAAGAGCAACAGAGTCCCTAAGTCTGTGTGCCAGCGAC	1500
OY	1501	tgtcttgaagggcaccagcagtggttagcgggttccatcacctgcgtcgttgtagtgtg	1560
Db	1501	TGTCTGAAGGGCACACGACAGTGTACGGTTCATCACTGCTGCTTGAAGTGTG	1560
OY	1561	ccctgtgagggtgagaccttccctcaacaagaagtgcacctacagatgcagacgttgtgag	1620
Db	1561	CCCTGTGGGCGCTGGAGCTTCCCTCAACAGAGAGCACTTACAGATGCCAGGCTTGTGG	1620
OY	1621	aaaaaagaatgaggaaccttaggggaagaacgaagcctgttccggcgacgttggtgttttg	1680
Db	1621	AAAAAGAATGAGGACCTGAGGGAAGCCAACTGCTTCCCGGCACTGTGTGTTTGG	1680
OY	1681	gctttagtgtagcaacaccttcttggtgtgtgtgtgcagcacaacgctgtgtgtgctgtg	1740
Db	1681	GCTTGTGCTGAGCACACCTCTGTGGTGTGCTGGCAGCTAACACGCTGTGCTGTGCTG	1740
OY	1741	ctgcttgggaatgtgtgacctgttctgctctggaactagaacccccgtgtgtgtgagtgcaaga	1800
Db	1741	CTGCTTGGGAATGTGTGCGCTGTGTGCTGGCACTTGAACACCCGTGTGTGAGTGCACGA	1800
OY	1801	ggggggccgcctgtccttcttctatgcgtgaggtgtccctctgacagcagtgtagtgagccctcat	1860
Db	1801	GGGGGGCGCGTGTCTTCTTATGCTGTGGCTCCCTGTGGCAGAGATAGTGGACGCTCTAT	1860
OY	1861	ggtctccttggggaaccacaagaagcctgtgtgtgtgtacagccctcttgcctt	1920
Db	1861	GGCTTCTTGTGGGGAACCCACAAGGCGCTGTGCTTCTACGGCAGGCGCTCTTGTGCCCTT	1920
OY	1921	ggttcaacaatcttcgtctcgtctgcgaagttgcgtatcttccaataatacatctc	1980
Db	1921	GGTTTCAACATCTTCTGTCTCTGTGCCTGTACAGTTGCTCATTTCCAACTATATCATCTTC	1980
OY	1981	aagtttccacaagaagtaccataatcttacaacagcctgtggttccaaaacacacggtgtgac	2040
Db	1981	AAGTTTCCACAAGAGTAACTACATTCTAACAGCGCTGGGTCCAAACACAGGTGTCTGC	2040
OY	2041	ctgttcttgatgatcagctcagcgcccgagctgtctatctgtctcaactgttgctgtgtgtg	2100
Db	2041	CTGTTTGTGATGATCACCTCAGCGGCCACAGCTCTATCTGTCTAACTGTGCTGTGTGTG	2100

QY	2101	tgagcccccagcctgctcgttaggaataacaagcgtcccccacatctggtgatgttaagtc	2160
Db	2101	tgaccccccagcctgctcgttaggaataacaagcgtcccccacatctggtgatgttaagtc	2160
QY	2161	acaagacaaatccctggtgctcactggtccttcactggtccttcatactggtccttcac	2220
Db	2161	acaagacaaatccctggtgctcactggtccttcactggtccttcatactggtccttcac	2220
QY	2221	atgctccttgctcagctgactggtgtaaggactgtcagaagaactacaagaagccaaa	2280
Db	2221	atgctccttgctcagctgactggtgtaaggactgtcagaagaactacaagaagccaaa	2280
QY	2281	tggtcactcctcagcctgctcttcaacttcgtgtccttgatcgtccttccaacgagcc	2340
Db	2281	tggtcactcctcagcctgctcttcaacttcgtgtccttgatcgtccttccaacgagcc	2340
QY	2341	agcgtctacgagcaggaagtactgctgctgcgcgcaacatgagctggtcgtgagagcctg	2400
Db	2341	agcgtctacgagcaggaagtactgctgctgcgcgcaacatgagctggtcgtgagagcctg	2400
QY	2401	agcagcggcttcgtggtgatttcttcgactaagtgactagtgatcctctgcgcgcagac	2460
Db	2401	agcagcggcttcgtggtgatttcttcgactaagtgactagtgatcctctgcgcgcagac	2460
QY	2461	ctcaacagcaacagagacttccaggctccatcagaactacaagaagcgtgagctcc	2520
Db	2461	ctcaacagcaacagagacttccaggctccatcagaactacaagaagcgtgagctcc	2520
QY	2521	acctga 2526	
Db	2521	acctga 2526	
RESULT 2	AX82900	2951 bp	DNA linear PAT 02-NOV-2001
LOCUS	AX82900	Sequence 9 from Patent WO0172842.	
DEFINITION	AX82900	Sequence 9 from Patent WO0172842.	
ACCESSION	AX82900.1	GI:1609872	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
BASE COUNT			
ORIGIN			
Query Match	100.0%;	Score 2526;	DB 6; Length 2951;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2526;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	atgcgtcctcgaagcgcgcgcctgctgagcctgcagagctcattcctcgtcgtgagcc	60
Db	232	atgcgtcctcgaagcgcgcgcctgctgagcctgcagagctcattcctcgtcgtgagcc	291
QY	61	tttgctcgcataagcaagagagctctcctcgaattcaacccctcccgagattactctcty	120
Db	292	tttgctcgcataagcaagagagctctcctcgaattcaacccctcccgagattactctcty	351
QY	121	gcagcgcgttccctcctcattctgctgctgtctgcagagtgagagcaagaacccgagtgacc	180
Db	352	gcagcgcgttccctcctcattctgctgctgtctgcagagtgagagcaagaacccgagtgacc	411

QY	181	ctgctgtgacaggtctctgtagcttcacatgacatgctgtacacccctcttcacggtcatgcg	240
Db	412	CTGTGTGACAGGCTCTTTACTCTTCATATGAGCAGTGGCTACACCTCTTCCAGGCTATGCGG	471
QY	241	cttgagggttgaaggagataaacaactccacagggccctctcccaatatcaccttgggtgac	300
Db	472	CTTGGGGTTGGAGGATTAACAACTCCACG6CCCTCTCTCCAAATCAACCTTG6GGTAC	531
QY	301	cagctgtatgtagtgtgttctgacctgcacaatgttatgcaagcttagagtctctcc	360
Db	532	CAGCTGATATAGTATGTGTCTTGTGACTGTGCCAAATGTGTATGCCAGCCTAGAGTGTCTCC	591
QY	361	ctgccagggaacaacacacatagagctcccaaggagacccctctccactatccctacggtg	420
Db	592	CTGCCAGGGGCAACACACATAGAGCTCCAAAGGAGACCTTCTCCATATTTCCCTACGGTG	651
QY	421	ctggcagtgatctggtgacctgacacacacacccgtgtctgccacaacagccgacctgtctgac	480
Db	652	CTGGCAGTGATTTGGGCGCTTGACACACACCAACCGTGTCTGCCACACAGCGCCCTGTGTAGC	711
QY	481	cccttccttggtgccacatgattagctatctgcccagcagcagagactcaacgttgaagcgg	540
Db	712	CCTTCTCTGGTGGCCCATGATTACTATGCGGCGCAGCAGGAGAACGCTCAGCGTGAAGGG	771
QY	541	cagatccctctctctctctgcacccatccccaatgacaagtaccagctgagacaatggtg	600
Db	772	CAGTATCCCTCTTCTCTGCGCACCATTCGCCAAATGACAACTACAGTATGAGACCATGATG	831
QY	601	ctgctgtctgcagaagttcctgggtgagacctggtatctctctgtgttgcaagcagtgaagactat	660
Db	832	CTGTGTCTGCGAAGAGTTCGGGGTGACCTGTGATCTCTGTGGTGAGAGATGACACTAT	891
QY	661	ggcagcactaagggtgtgcagagacctggaagcaacagccacctgctcagggatctgcatgct	720
Db	892	GGCGAGCTAGGGGTGCGAGCAGCAGCACTGAGAAACAGCCACTGTGTACAGGATCTGCATTCT	951
QY	721	ttcaagagacacatgcccctctctctgccaggctgggcgatgtagaagatgcaatgacctcatg	780
Db	952	TTCAAGGACATCATGACCCCTTCTCTGCCAGGTGGCGATGAGAGATCAGAGCTCCATG	1011
QY	781	cgcacacttgccccagagccggggggccacccgtcgtgtgtgttttctcagccgggaattggcc	840
Db	1012	CGCACCTGGGCCCCAGGCGCGGGGCGACCGTGTGTGTTTTCAGCGCGGAGTGGGCC	1071
QY	841	agggtgttttttcagatctccgtgctgcctgcacaaacctgacgtgcgaagtcttggtctcccta	900
Db	1072	AGGGTGTTTTTCAGATCCGTGTGCTGCTGACCAACCTGACTGCGAAGTGTGGTCCCTCA	1131
QY	901	gaagcccttgggccctctccaaagcacatcactggtgggtgcgccgggatatccagcgaattggatg	960
Db	1132	GAAAGCTGGGGCCCTCTCCAGGACATACACTGGGGGTGCCGGGATCCAGGCGCATTTGGATG	1191
QY	961	gtgctgtggcgttggtccatccagaagaggtctgtccctggtccctaaaggctttgaaagaagcc	1020
Db	1192	GTGTGTGGCGCGGCGCATCCGAAGAGGGCTGTCTCCGTTCAMAGCGCTTTGAAACAAACC	1251
QY	1021	tatccccggggaagaagaaggagcccttaaggccttgcacaagaaggtctcctgtgtgagcagc	1080
Db	1252	TATGCCCCGGGCAACAGAGAGGAGGCCCTTAGGCTTGGCCACAAAGGCTCTCTGTGACAGCAC	1311
QY	1081	aatcacgtcttgcaagaagaatgcacaagtttcaatgtrcacacaagatgcccaagctcaaaagcc	1140
Db	1312	AATCAGGCTGTGCAGAGAAATGATCCCAAGCTTTCATRCACACAGATGCCCAAGGTCAAAGCC	1371
QY	1141	ttctccatgagttctgtcctacaacgaataccgggtctgtatgtggttggtccatggtctc	1200
Db	1372	TTTCTCCATGAGATTCTGTGCTCAACAACGCAATCCGGGCTGTGTATGCGGTGCGCATGTGCTC	1431
QY	1201	caccagactccctgggctctgtgctcttgtagagcttfttccaaaggccaggtcttaacccctggag	1260
Db	1432	CACCAAGCTCTTGGGCTGTGCTCTGAGAGCTGTGTTCCAGGAGGCGAGATCTACCCCTGGGAG	1491

QY 1261 ctttggagcaagatccacaagtgatcttctctacacagaagacgtgagcgtttaa 1320
 |||||||
 Db 1492 cttttggacagatccacaagtgatcttctctacacagaagacgtgagcgtttaa 1551
 |||||||
 QY 1321 gacaaagagatccctcctagtgatacaataattgcttggaactggaatggccaa 1380
 |||||||
 Db 1552 gacaaagagatccctcctagtgatacaataattgcttggaactggaatggccaa 1611
 |||||||
 QY 1381 tggacctagcgtctcctggttccacatgcttcacagttcagttcaataatggag 1440
 |||||||
 Db 1612 tggacctagcgtctcctggttccacatgcttcacagttcagttcaataatggag 1671
 |||||||
 QY 1441 accaaatccagtgagcaaggaagaaacacagtgagcctaagtcgtgttccagcag 1500
 |||||||
 Db 1672 accaaatccagtgagcaaggaagaaacacagtgagcctaagtcgtgttccagcag 1731
 |||||||
 QY 1501 tgtcttgaaggcaccaggaagtgtgttaagggtttccacactgctgcttgaagtgt 1560
 |||||||
 Db 1732 tgtcttgaaggcaccaggaagtgtgttaagggtttccacactgctgcttgaagtgt 1791
 |||||||
 QY 1561 cctctgagagcttgagaccttcccaacaagatgacctctacagatgcaagccttgag 1620
 |||||||
 Db 1792 cctctgagagcttgagaccttcccaacaagatgacctctacagatgcaagccttgag 1851
 |||||||
 QY 1621 aaagaagatgggacactgaagggaagcagacactgctcccgagcagctgtgtgtttg 1680
 |||||||
 Db 1852 aaagaagatgggacactgaagggaagcagacactgctcccgagcagctgtgtgtttg 1911
 |||||||
 QY 1681 gcttgagcagacacacacttggtgtgtctgagcagcttaacagcgttgctgctgctg 1740
 |||||||
 Db 1912 gcttgagcagacacacacttggtgtgtctgagcagcttaacagcgttgctgctgctg 1971
 |||||||
 QY 1741 ctgcttgagacgtgagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1800
 |||||||
 Db 1972 ctgcttgagacgtgagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2031
 |||||||
 QY 1801 gggggcgagcgt 1860
 |||||||
 Db 2032 gggggcgagcgt 2091
 |||||||
 QY 1861 ggcctcttgagggaacccaagagcgtgctgtgtgtgtgtgtgtgtgtgtgtgtgt 1920
 |||||||
 Db 2092 ggcctcttgagggaacccaagagcgtgctgtgtgtgtgtgtgtgtgtgtgtgtgt 2151
 |||||||
 QY 1921 ggttccacatctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1980
 |||||||
 Db 2152 ggttccacatctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2211
 |||||||
 QY 1981 aagtttccacaagaagttactatcttccacagcgtgtgtgtgtgtgtgtgtgtgt 2040
 |||||||
 Db 2212 aagtttccacaagaagttactatcttccacagcgtgtgtgtgtgtgtgtgtgtgt 2271
 |||||||
 QY 2041 ctgttgatgatcagctcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
 |||||||
 Db 2272 ctgttgatgatcagctcagcagcagcagcagcagcagcagcagcagcagcagcagc 2331
 |||||||
 QY 2101 tggacccaactgctgt 2160
 |||||||
 Db 2332 tggacccaactgctgt 2391
 |||||||
 QY 2161 aagaagacaaactcctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2220
 |||||||
 Db 2392 aagaagacaaactcctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2451
 |||||||
 QY 2221 agtgacctgctgcaactgcaactgcaactgcaactgcaactgcaactgcaactgca 2280
 |||||||
 Db 2452 agtgacctgctgcaactgcaactgcaactgcaactgcaactgcaactgcaactgca 2511
 |||||||
 QY 2281 tggatcacttaagcctgtcttcaactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2340
 |||||||
 Db 2512 tggatcacttaagcctgtcttcaactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2571
 |||||||
 QY 2341 agcgtctagaagcaggaactgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2400
 |||||||

Db 2572 AGGCTTACAGACGGCAAGTACTGCTGGGCGCCACATGATGCTGGGCTGACACCTG 2631
 |||||||
 QY 2401 agcagcgtctgt 2460
 |||||||
 Db 2632 AGCAGCGGCTGCGTGGGATTTTGTGCTTAAGTGTACTGTGATCTGTGCGCCAGAC 2691
 |||||||
 QY 2461 ctcaacagcacaagacacttccagcctccatccatcaagactacagaagcgtgagctc 2520
 |||||||
 Db 2692 ctcaacagcacaagacacttccagcctccatccatcaagactacagaagcgtgagctc 2751
 |||||||
 QY 2521 acctga 2526
 |||||||
 Db 2752 ACCTGA 2757
 |||||||
 RESULT 3
 AX354027 2759 bp DNA linear PAT 06-FEB-2002
 LOCUS Sequence 13 from Patent WO0203848.
 DEFINITION AX354027
 ACCESSION AX354027
 VERSION AX354027.1 GI:16618964
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Naito A.T.
 TITLE Method for opening the blood-brain barrier
 JOURNAL Patent: WO 0203848-A 13 17-JAN-2002;
 NAITO, Albert T. (US)
 FEATURES
 source Location/Qualifiers
 1..2759
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7475203CB1"
 BASE COUNT 523 a 828 c 782 g 626 t
 ORIGIN
 Query Match 91.9%; Score 2322; DB 6; Length 2759;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2522; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 atgtgtctctgcaagcgtcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 60
 |||||||
 Db 53 ATGCTGCTGTGCAAGGCTGCGTGGGCTGACGCTTCTCATTTCTGCTGCGGCC 112
 |||||||
 QY 61 ttgtctgcatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 120
 |||||||
 Db 113 TTGCGTGCATAGCAGCGGAGTCTTCTCGACTTCAACCTCCCGAGATTACCTCTG 172
 |||||||
 QY 121 gcaagccttctcctctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 180
 |||||||
 Db 173 GCAGGCTGTGCTCTCTCCATTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 232
 |||||||
 QY 181 ctgttgcaagcgtctgt 240
 |||||||
 Db 233 CTGTGTGACAGGCTGTGTGCTTCAATGAGCATGGCTTACACACTCTTCCAGGCTATG 292
 |||||||
 QY 241 ctgttggttgagagataaacaactccacagcgtctgtgtgtgtgtgtgtgtgtgtgt 300
 |||||||
 Db 293 CTTGGGCTTGAGAGATTAACAACATCCACGCGCTCTGCTGCCAATCACCCGTGGGTAC 352
 |||||||
 QY 301 cagctgtatgagt 360
 |||||||
 Db 353 CAGCTGTATGATGT 412
 |||||||
 QY 361 ctgccagagcaaacacacatagagctccacagagacacttctcactatcccttaggtgt 420
 |||||||
 Db 413 CTGCCAGGGCACCACCATAGAGCTCCAAAGGAGACCTTCTCAGCTATTTCCCTAGGCTG 472
 |||||||

421 ctgagcagtgattggcctgacagcaacacgctgtctgcacccacagccgctctgtgagc 480
|||||
Db 473 CTGGAGATGATTGGGCTGTACAGCAACCTGCTGCCACACAGCCCTGCTGAGC 532
Oy 481 ccttccctgtgtgccaatgattagctatagtcgcgcagcagcagcagcgtctcagcgttgaacg 540
|||||
Db 533 CTTTTCCTGTGGTCCCATGTATGTAGTATGGGGCCAGCAGCAGCAGCCTCAGCTGAAGCGG 592
Oy 541 cagttccctcttctctctgcgcacacaccccaatgacaaagtaacaggttgagacacatgtgt 600
|||||
Db 593 CAGTATCCCTCTTCTCTGGGACCAATCCCAATGACAAATACAGTGGAGACCAATGTG 652
Oy 601 ctgcctgtgcagaaagctcgcgtgtgacacgtatctctctgtgttgagcagcagtaacacat 660
|||||
Db 653 CTGCTGTGCAGAAATTTGGGTGGACCTGTGATCTCTGTGTGGAGCAGCTACGACTAT 712
Oy 661 gggcagcctaggggtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
|||||
Db 713 GGGCAGCTAGGGGTGACAGGACCTGAGAAACAGGCCACAGTGTGAGGGGATCTGCAATTGCT 772
Oy 721 ttcaagagacatcatgctctctctctgcacaggtgtgagcagatgagagatgagcagctcatg 780
|||||
Db 773 TTCAAGGACATCATCCCTTCTCTCCAGGTGGGCGATGAGAGATGACAGTGCCTCATG 832
Oy 781 cgcacacctgagccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 840
|||||
Db 833 CGCCACCTGGGCCACGGCCGGGSCACCGTGTGTGTTTTCACGCCGCGAGTGTGGC 892
Oy 841 aggtgttcttctgagtcggtgtgtctgacaaactgtacgtgcagcagcagcagcagcagcagc 900
|||||
Db 893 AGGGGTGTTTTCAGTCCGTGTGTGTGACCAACCTGACAGCAGCAAGTGTGGGTCCCTCA 952
Oy 901 gaagcctgtggccctctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
|||||
Db 953 GAGGCTTGAGCCCTTTCAGGACATCACTGGGTGTGCCGGATCCAGGCAATGTGGATG 1012
Oy 961 gtctgtgagcgtgtgcatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1020
|||||
Db 1013 GTGCTGGGGTGGCCATCCAGAAAGGGGTGTCCCTGGCTGAAGGCGTTTGAAGAGCC 1072
Oy 1021 tatgcgcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
|||||
Db 1073 TATGCCCGGGGAGACAGAAAGGCCCTTGAAGCCTTGCACAAAGGCTCCGTGTGACAGCAC 1132
Oy 1081 aatcagcctctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
|||||
Db 1133 AATCAGCTCTGCAGAGATGCCAAAGCTTTCATGTGCAACAGATGCCAAAGCTCAAAAGCC 1192
Oy 1141 ttctccatgagctctgcctacacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1200
|||||
Db 1193 TTCTCCATGAGATTTCGCTACAAAGCATACCGGGGTGTATGGCGTGGCCATGGCCTTC 1252
Oy 1201 caacagcctcgtgagcgtgtgcctctgagagctgttccaggggcagagctacacccctgcag 1260
|||||
Db 1253 CACAGCTCTGCGGTGTGCTCTGTGAGCTTTCAGAGGGCGGAGTCAACCCCTGGGAG 1312
Oy 1261 cttttggagcagacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1320
|||||
Db 1313 CTTTGGAGCAGATCCACAGGTCATTTCTTTCACACAGAGACCTGTGGCGTTTAAAT 1372
Oy 1321 gacacacagagatccctcagtagctataacataatgtccctgtgagcagcagcagcagcagcag 1380
|||||
Db 1373 GACACAGAGATCCCTCAGTAGTATATAATTTGCTGGAGCTGGAATGAGACCCAAAG 1432
Oy 1381 tggaccttcagcgtctcgtgtctcctccacatgtgttccagctcagcagcagcagcagcagcag 1440
|||||
Db 1433 TGGACCTTTCACGCTCTGTGTCTTCACATGTGTCCAGTTCACCTAAACATAATATGAG 1492
Oy 1441 accaaatccagctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500
|||||
Db 1493 ACCAAATTCAGTGGCAGCGAAGACACAGGTGCTTAAAGTGTGTGTCCAGCGAGC 1552
Oy 1501 tgtcttgagggcaccagcagctgtgtacaggtttccatcactgtgtcttgagctgtgt 1560

|||||
Db 1553 TGTCTTGAAGGGCACACAGGAGTGTAGGGTTTCATCTACCTGCTTGTAGTGTGTG 1612
Oy 1561 cccctgtgtgtgtgtgagccttccctcaacagagtgacctctacagatgacagcctgtgtgtg 1620
|||||
Db 1613 CCTGTGGGGCTGGGACCTTCTCAACAAGAGTGTACCTTACAGATGCCAGCTTGTGGG 1672
Oy 1621 aaagaagagtgagccttgaaggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680
|||||
Db 1673 AAAGAAAGTGGCACCTAGGGAACCCAGACCTGCTTCCGGCAGCTGTGTGTGTG 1732
Oy 1681 gcttctgtgagcagacacccctctgt 1740
|||||
Db 1733 GCTTGTGAGACACACCTTGT 1792
Oy 1741 ctgcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800
|||||
Db 1793 CTGCTGGGAGTGTGGCTGT 1852
Oy 1801 gggggcgcgt 1860
|||||
Db 1853 GGGGGCCTGT 1912
Oy 1861 ggtcttctgtgggaaccccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1920
|||||
Db 1913 GGTCTTGTGGGGAACCCACAAAGGCGCTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1972
Oy 1921 ggttcaacatctctctgt 1980
|||||
Db 1973 GGTTCACCATCTTCTGT 2032
Oy 1981 aagtttccacaaaggtactacatctacacagcagcagcagcagcagcagcagcagcagcagcagc 2040
|||||
Db 2033 AAGTTTCCACAAAGTACTTACATTTTACACAGCTGTGGTCCAAACACAGGTGTGGC 2092
Oy 2041 ctgttctgtatgatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
|||||
Db 2093 CTGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2152
Oy 2101 tggaccccaactgt 2160
|||||
Db 2153 TGGACCCCACTCTGT 2212
Oy 2161 acagagac 2220
|||||
Db 2213 ACAGAGACCAACTCCCTGGGCTTCAATCTGAGCTTCTCTACAAATGAGCCTCTCTTCATTC 2272
Oy 2221 agtgccttctgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2280
|||||
Db 2273 AGTGCCTTGTGCTGAGCTAGCTAGGTTAGGACTTGCAGAGAACTAACAGAGGCCA 2332
Oy 2281 tgytcaaccttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2340
|||||
Db 2333 TGTGTACCTTCAAGCTGT 2392
Oy 2341 agcgtctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400
|||||
Db 2393 AGCGTCTACAGAGGCAAGTACTGCTGT 2452
Oy 2401 agc 2460
|||||
Db 2453 AGCAGGCGCTGT 2512
Oy 2461 ctcaacagc 2520
|||||
Db 2513 CTCAACAGCAGAGACTTCCAGGCTTCAATTCAGGACTACAGAGAGGCGTGGCTTC 2572
Oy 2521 acctga 2526
|||||
Db 2573 ACCTGA 2578

RESULT 4

[illegible]

Query Match	90.7%	Score 2292	DB 6	Length 2292
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2292	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 235	atgcgctctggggttctgagagataaacacatccagccctctgcccacataccctg	294		
Db 1	ATGCGGCTTGGGGTTGAGAGATTAACAACACTCCAGCGCCTCTCCACACTACCTG	60		
QY 295	gggtacagctgtcatgatgtgttctgttctgtacctctgccaatgtgatcatcagcttgaagt	354		
Db 61	GGGTACCAAGCTGTATGATGATGTGTGTGTGACTCTGCCAATGTTGATGCACCTGTAGAGT	120		
QY 355	ctctcccccgcgaaggcaacacacaaatagaagctccaaagagacctcttcacattccct	414		
Db 121	CTCTCCCTGCGCAAGGCAACACCACATTAAGCTCCCAAGAGACCTTCTCCACTATTCCCT	180		
QY 415	acggtctgcgagtgatcttgggctctacagcaacacacgtctctgcacacaaacgcgcctg	474		
Db 181	ACGGTGTGGGCAAGTATTTGGGCTCTACAGCACCAACCGTGTGCGCACCAACGCCCTG	240		
QY 475	ctgagcccttccctgtgtgcccatagtatagctatgctgcgccaagcagagagctcagctg	534		
Db 241	CTGAGCCCTTTCCTGCTGATGCTATGCTATGCGGCGACGAGCGAGACGCTCAGCGTG	300		
QY 535	aagcgagagatctccctcttctctgtgcgcaaatccccaatgaataagttacagttgaagcc	594		
Db 301	AAGCGGCAAGTATCCCTCTTCTCGGCAACCACTCCCAATGAACATACCAGGTGAGACC	360		
QY 595	atggtctgcgtctgcataaagttctgggtgcagactgatactctctgttgcagcagatgac	654		
Db 361	ATGGTGTGCTGTCTGCAAAAGTTCGGGTTGGAGACTTGATCTCTCTGTGGCGACAGTGAC	420		
QY 655	gactatgggcagctgaggggtgcagagcacttggaaacacagcccaactgctcaggggatactgc	714		
Db 421	GACTATGGGCAAGCTATGGGGTGTGACGACTGTGAGAACCAAGCCATGTGTAGGGGATGTGC	480		
QY 715	atctgcttcaagaagataatctgcctctctctgcgccaaggttgggcgatagaagaatctcaatgc	774		
Db 481	ATTGCTTTCAAGGACATATGCGCTTCTCTGCGCCAGGTGGGGATGAGAGGATGCAATGTC	540		
QY 775	ctcaatggcacaactggcccaagccgggcccacacgctcgtgtgttttcttcagccggag	834		
Db 541	CTCATGGCCCACTGTGGCCCAAGGCCGGGCCACACGTGTGTGTTTTCACAGCGGCGAG	600		
QY 835	ttggccagagtgatttcttcagatcgttggtctgcagcaaaccttgcagcgaagttctgggtc	894		
Db 601	TTGGCCAGAGGTGTTTTCAGATGCGGTGATGCTGACCAACACTGTACTGTGCGAAGTGTGGGT	660		
QY 895	gacctagaagcttggccctctccagacacatacttgggtgtgccggatccagcgact	954		

Db	661	GCCTAGAAAGCTGGGGCCCTCTCCAGAGCACATCACTGGGGTGGCCCGGATCCAGGCCATT	720
Oy	955	ggagatggtgtgtggtgcgttggccatccagaaagagctgttccctgtgacctgaagcggtttga	1014
Db	721	GGGATGGTGTCTGGGGGTGGCCATCCAGAAAGAGGGGTGTCTCCGGCTGAAGGCGTTTGA	780
Oy	1015	gaagccttctcccggtgagacaaagagagggcccttaagccttgcacaagaggtccctgtgc	1074
Db	781	GAAGGCTATGGCCCGGGGACACAAAGGAGGCCCTTAGGCCCTTGGCACAGGGGCTCTGTGTGC	840
Oy	1075	agcagcaatcagctctgtcagagaatgcccagcttcatatgcaacacagatgtcccaagctc	1134
Db	841	AGCACATATCACTGCTGGAGAGATGGCAAGCTTTCATGRCAACACATGCGCCAAAGCTC	900
Oy	1135	aaagccttcctcctatgaatgtctgtctactacaacgcataccgggtcggtatgtatgtgtgccc	1194
Db	901	AAAGCTTCTCCATGATTCATGTTCTCCCTCAACAACGATACCGGGGCTGTATCCGGTGGCCAT	960
Oy	1195	ggccctcaacagctcctctgggtgtgtgcctcttgagcttgttccagaggccgaatctacccc	1254
Db	961	GGCCTCCACAGACTCCTGGGCTGTGTGCTCTGTGAGCTTGTCCAGGGGGCCGAGCTTACCCC	1020
Oy	1255	tggcagcttttggagcagatccacacaagtgcatcttccctctacacaaggacatctgtgcg	1314
Db	1021	TGGCAGCTTTTGGACAGATCCACAAAGGAGCTTTCCTTCACAGAGCACACTGTGGGG	1080
Oy	1315	tttaatgcaaaagagatccctctagtgatctaacaatatatgtccctggagctggaaatga	1374
Db	1081	TTTAAATGACAAAGAGATCCCTCCATGATACATAATATATGCTTGAGACTGGAAATGGA	1140
Oy	1375	cccaagtgagccttcaacaggtccctgtgttccctcaaatgtgtctccagttcagatlaaacata	1434
Db	1141	CCCAAGTGAGACTTACAGGCTCTCGGTTCCACATGCTCTCCAGTTCAAGTAAACATA	1200
Oy	1435	aatgagaccaaaatccagatgtgagacggaaagacaaaccgggttccctaaagtcttgtgttcc	1494
Db	1201	AATGAGACCAAAATCCAGTGGGACGGAAGAAAGCAACAGGTGCTTAAGTCTGTGTGTCC	1260
Oy	1495	agcagactctcttgaagggcaaccagagatgtgttactaggttttccatcaactgtcgtcttga	1554
Db	1261	AGCGACTGTCTGAAGGGGACCAAGGAGTGTACAGGGTTTCCATACATGCTGTGTGAG	1320
Oy	1555	tggtgtccctgtgtgggtcttgggaacttctctaaacaagatgtgaactctacagatgcaagcct	1614
Db	1321	TGTGTGCCCTGTGGGGCTGGGACCTTCCCAACAAGAGTACCTTACAGAGCCAGCCT	1380
Oy	1615	tgttggaaagaaagtgtgagcaacctgagaaagcaagacctgttcccggtgcaactgtgtg	1674
Db	1381	TGTGGGAAAGAAATGTGGGCACTGAGGAAACCAAGACTGTTCCCGGCACTGTGTGTG	1440
Oy	1675	tttttgcttgcgtgtgagcaacactcttgggtgtgtgtgtgtgcaagctaaacgctgtctgt	1734
Db	1441	TTTTTGGCTTTGGCGTAGAGCACACTCTTGTGGTGTCTGTGGAGCTTAACAGCTGTGCTG	1500
Oy	1735	ctgctgcgtcttggagactgtgcgtgtgttctgtccttgcaactagaacaccttgttgaag	1794
Db	1501	CTGCTGTCTTGGGAGCTGTGCTGTGTGTGTGCTGTGCACTTGAACACCCCTGTGTGTGAG	1560
Oy	1795	tcaagaagggagcgcgctgtcttcttctatgtctgggtccctctggccaagtagtggcagc	1854
Db	1561	TCAGCAGGGGGCGCGCTGTCTTCTTATGTGGGCTCCCTGGCAGAGGTGTGTGCAAGC	1620
Oy	1855	ctctatgtctcttctggggaacccaagaagctgtgtgtctgtctagccagagcctcttt	1914
Db	1621	CTCTATGGCTTCTTGGGGAAACCCAAAGGCTGTGGTGTCTTACGCGCAGGCCCTCTT	1680
Oy	1915	ggccttggtttacaacatcttccctgcacbcctggaagttcgctcatctccaataatcatc	1974
Db	1681	GCCCTTGGTTTACCATCTCTCTGTCTGTCTGTGACAGTTCGGTCAATTCAACTAATATCATC	1740
Oy	1975	atcttcaagtttccacaaggttacctatcttaccacagccttgggtccaaaacaggt	2034

Db	1741	ATCTTCAGTGTTCACCAAGGTACCTACATCTTCACACAGCCTGGGTCCAAACACAGGT	1800
QY	2035	gctggcctgttttgatgatcatcagctcagcgcgcagctgtctatctgtctaaactgtgctg	2094
Db	1801	GCTGGCCTGTTTGTGATGATCAGCTCAGCAGCGGCCAGCGCTTATCGTCTACTGGCTG	1860
QY	2095	gtgagtgtgagccccaactgctgtctatgggaataccaagcgcttccccatctggtgatgctt	2154
Db	1861	GTGGTGTGGACCCCACTGCTGTGGGAAATACAGCGCTTCCCACTGTGGTGAAGCTT	1920
QY	2155	gagtgcagagagacccaactccctggctgtctcatatcgtgcctctcttaaatgacctcctc	2214
Db	1921	GAGTGCACAGAGACCAACTCCCTGGGCTTTCATCTAGCTGGCTTCTCTTACATAGGCTCTCTC	1980
QY	2215	tcacatcagttgccttctgtcctgcagctcagcttggttaaggacttgcagagaaactacaagag	2274
Db	1981	TCCATCAGTGCCTTTGGCTGCAAGTACCGTGGGTAAAGACTTCCAGAGAACTACAAACAG	2040
QY	2275	gccaaatggttcaacttcaagctgtctcttcaacttcgtgtccctgtatcgacttctcac	2334
Db	2041	GCCAAATGTGTACACTTTCAGCTGCTGCTTTCAACTTCGTGCTCTGGATGCGCTTCTTCAC	2100
QY	2335	acggcgcagcgtctaaagacggaagaaactgcctgtgggccaataatagtgtggtgtgac	2394
Db	2101	ACGGCCACAGCTTACGACGGCAAGTACCTGCTTGCGGCCAACATATATGGCTGGGTGAC	2160
QY	2395	agcctgagcagcgcgttcgtgtggtattcttcgcctaaagtgtctaagtatccctgcgcg	2454
Db	2161	AGCCGAGACAGCGGCTTGGTGGGATTTTTCGCTTAAGTGTACTGATCTCTCTGCGCC	2220
QY	2455	ccaagacctaacaagacagagaaacttccagcgtccatccatccaagacttaacagagagcttgc	2514
Db	2221	CCAGACCTCAACAGACAGACAGACATTCAGAGGCTCCATTCAGAGACTACACGAGGGCGTGC	2280
QY	2515	ggctccacctga 2526	
Db	2281	GGCTCCACCTGA 2292	

RESULT	5				
AX282896					
LOCUS	AX282896	1101 bp	DNA	linear	PAT 02-NOV-2001
DEFINITION	Sequence 5 from Patent WO0172842.				
ACCESSION	AX282896				
VERSION	AX282896.1	GI:16609870			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (sites)				
AUTHORS	Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,				
	Abidin,A., Zambrowicz,B. and Sands,A.T.				
TITLE	Novel human 7tm proteins and polynucleotides encoding the same				
JOURNAL	Patent: WO 0172842-A 5 04-Oct-2001;				
	Lexicon Genetics Incorporated (US)				
FEATURES	Location/Qualifiers				
source	1..1101				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	219 a	308 c	323 g	250 t	1 others
ORIGIN					

Query Match	39.5%	Score 999	DB 6	length 1101
Best Local Similarity	100.0%	Pred. No. 0		
Matches 999	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Oy	555	atgtgtgcctgcctgcgcacaaattctgggtgtggaacctgcgatactctctctgtttggcagcagtgac	654
Db	1	atggctgcctgcctgcgcacaaattctgggtgtggaacctgcgatactctctctgtttggcagcagtgac	60
Oy	655	gactatggcagcctctgggtgtgcagcgaactgtgaagaaccagcgcgaactgtgtcagcaggaattcgc	714

Db	61	GACTATGGGACACTAGGGGTGCAAGGACATGGAACCAAGCCACTGGTATAGGGGAACTTC	120
QY	715	attgctttcaagaagatcatctgcctctctctgcgcaagtgaggatgagagatctgaatgc	774
Db	121	ATTGCTTTCAAGAGACATCATGCCCTTCTCTGCCCCAGGTGGGGATGAGAGATGCAAGTCC	180
QY	775	ctcatctgagccactctgcccaggccgggagccacgctctgattgttttttccagccggcag	834
Db	181	CTCATGAGCCGACCTTGAGCCAGGCCGGGGCCACGCTGCTGTGTTTTCAGACGGGACG	240
QY	835	ttggcccaagggttttttcgagtcggtgtgtgtcgtgacccaactctggaaggtgtgtgtc	894
Db	241	TTGGCCGAGGGTGTTTTTCAGTGCCTGAGTGCAGCCAACTCAGTGGCAAGGTGTGGGTC	300
QY	895	gctcagaagagccttgaggccctctccaggacacactgaggggtgcccggatcccaagccat	954
Db	301	GCCCTAGAAAGCTGGGGCTCTCTCCAGGCAACATCTAGGGGTCCCGGGATCCAGGCAAT	360
QY	955	ggagatgtgtctggtgagctggcctccacagaagaggtctgtccctgcgctgaaggcgttga	1014
Db	361	GGGATGTGTCGTGGGCGTGGCCATCCAGAAAGAGGGGTGTCCTCGGCTGAAGCGTTTGA	420
QY	1015	gaagcctatgcccgggacagaagaagagcccttaaggccttgcacaaggctcctgtgtgc	1074
Db	421	GAAGCCTATGCCCCGGGCAACAAAGGAGGCCCTTAGGCCCTTGGCACAAAGGCTCTGTGTGC	480
QY	1075	agcagaacatcagctctgcagagaatgagcaagtttcattatgcacaagagatgcccgaagc	1134
Db	481	ACACAGCAATCACTCTGCGAGAGAAATGCCAAGCTTTTCATGATGACACACGATGCCAAGCTC	540
QY	1135	aaagcctctccatgattctctcctacaagcagatcacccgggctgtgtatcggtggccat	1194
Db	541	AAAGCCTTCTCATGAGTCTCTGCTACAAAGGATACCGGCTGTGTATCGGTGGCCAT	600
QY	1195	ggcctccaccagctcctgtggctgtgtcctcttgagcttgttccaggggccgagttctacccc	1254
Db	601	GGCCTTCACACACTCTGGGCTGTGCTCTGTGAGCTTGTTCAGAGGGCCGAGTCAACCC	660
QY	1255	tggcagcttttgagcagatccacaagtgatcttccctctctacacaaggaactctgtgcg	1314
Db	661	TGGCAGCTTTTGAGACAGTCCACAAAGGTGCTTTCCTTCAACAAAGGACACTGTGGCG	720
QY	1315	tttaatgacaagaagatccctcctagtagctatataatattgctctggagccggaatgga	1374
Db	721	TTTATATGACAAGAGATCTCCTCTAGTAGCTATTAACATATATTGGCTGGGACATGGAATGA	780
QY	1375	ccaagtggacctgaagctcgtctggttccctacatgtgtctcaagttcaagttaacata	1434
Db	781	CCCAAGTGACTTTACGGTCTCTCGGTTCTCTCAATGATGTCTCCAGTTCCAGTTAAACATA	840
QY	1435	aatgagacaaaatccagctgagcaggaagaacaaccagtgacttaagtctgtgtctcc	1494
Db	841	AATGAGACCAAAATCCAGTGGGACGGAAGAAAGACAAACAGTGCTTAAGTGTGTGTCC	900
QY	1495	agcgactgtcttgaaaggagcacagcgagtgtctacgggtttccatcatcgtctgtttgag	1554
Db	901	AGCGACTGTCTTGAAGGGGACACGAGCGAGTGGTTACGGGGTTTCCATCACGCTGCTTTGAG	960
QY	1555	tgtgtgcccctgtgggtctgggagccttctccaacaagat	1593
Db	961	TGTGTGCCCTGTGGGGCTGGGACCTTTCCTCAACAAGAGT	999

RESULT	6
AL591866	
LOCUS	AL591866 102539 bp DNA linear PRI 25-OCT-2001
DEFINITION	Human DNA sequence from clone RP11-58A11 on chromosome 1, complete
ACCESSION	AL591866
VERSION	AL591866
KEYWORDS	AL591866.13 GI:16501194
SOURCE	HTG.
	human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 102539)
 AUTHORS Wallis, J.
 TITLE Direct Submission
 JOURNAL Submitted (25-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On Oct 26, 2001 this sequence version replaced gi:15022299.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw:
 SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr1
 RP11-58A11 is from the library RPEC1-11.1 constructed by the group
 of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-58A11. It may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true right end of clone RP11-58A11 is at 102539 in this
 sequence. The true left end of clone RP11-239P22 is at 13247 in
 this sequence. The true right end of clone RP4-650H14 is at 2000 in
 this sequence.

FEATURES
 source
 1..102539
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-58A11"
 /clone_11b="RPEC1-11.1"
 582..618
 /note="Sequence from overlapping clone RP4-650H14
 (AL158217). Assembly confirmed by restriction digest."
 1452..1473
 /note="Sequence from overlapping clone RP4-650H14
 (AL158217). Assembly confirmed by restriction digest."
 48541..48603
 /note="Sequence from overlapping clone RP11-239P22
 (AL159177). Assembly confirmed by restriction digest."
 69007..69267
 /note="Sequence from overlapping clone RP11-239P22
 (AL159177). Assembly confirmed by restriction digest."
 100154..100179
 /note="Sequence from overlapping clone RP11-239P22
 (AL159177). Assembly confirmed by restriction digest."
 BASE COUNT 25289 a 26768 c 27360 g 23122 t
 ORIGIN

Query Match 34.9%; Score 882; DB 9; Length 102539;
 Best local Similarity 99.9%; Pred. No. 0;
 Matches 932; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1594 gacctacagatgccagctctgttggaagaagatgggcaactgaggaagccagacc 1653
 |||||||
 Db 75842 GACCTCTACAGATGCCAGCCTGTGGGAAGAAGACTGGGCACTTAGGGAAGCCAGACC 75901
 QY 1654 tgcctccgcgcgaactgtgtgtgtttgtgcttgcgtgaacaaacctctgggtgcctg 1713
 |||||||
 Db 75902 TGCTTCCCGCGCACTGTGTGTTTTGGCTTGGCGGACACACCTCTGGGTGCTGCG 75961
 QY 1714 ggcgtcaaacgctgtctgtctgtcgtcgtcttgaggacgcctgagctgttgcctgac 1773
 |||||||
 Db 75962 GCAGCTAACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 76021
 QY 1774 ctagaacacctgtgtgtgtgaagtaagaaggggagccgctgtgtcttctatgtcgtgctc 1833
 |||||||
 Db 76022 CTAGACACCCCTGTGTGTGAGGTGACGAGGGGGCGCGCTGCTTGTATGTGCGGCTCC 76081
 QY 1834 ctggcgcgcggtgtgtgtgcagcctctatgtcttctttggggaacccaagagctgtgc 1893
 |||||||
 Db 76082 CTGGCAGCAGGTAGTGGCAGCCCTGTATGCTTGTGGGGAACCCAGAGGCTGCTGC 76141
 QY 1894 ttgctacgcagagccctcttctgtcctgtgtttacaccatctcctgcctgcctgacagt 1953
 |||||||
 Db 76142 TTGCTACGCCAGGCCCTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 76201
 QY 1954 cgcctacccaactaatcaltcaltcaltcaltcaltcaltcaltcaltcaltcaltcaltc 2013
 |||||||
 Db 76202 CGCTCATTCCAACTATATCATCTCTTCAAGTTTCCACCAAGATACCTCATCTCTACAC 76261
 QY 2014 gctctgtctcaaacacagagtgctgtgcctgtgtgtgtatgaatgaagctcaagcgcagctg 2073
 |||||||
 Db 76262 GCTGTGGTCCAAAACACAGGCTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 76321
 QY 2074 ctatctgtctaaactgt 2133
 |||||||
 Db 76322 CTATCTGTCTAACTTGTGGT 76381
 QY 2134 ttcctccatctgt 2193
 |||||||
 Db 76382 TTCCCTCCATCTGT 76441
 QY 2194 ttcctctacaatgtgctctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2253
 |||||||
 Db 76442 TTCTCTTCAATGCT 76501
 QY 2254 ttgccagagaactacaacagagcgaatggtacaccttcaagcctgtccttcaactcgtg 2313
 |||||||
 Db 76502 TTGCCAGAGAACTACAACGAGGCGCAATGTGTACCTTCAAGCTGTCTTCAACTTCGTG 76561
 QY 2314 tctctgatgcctctcttcaacagcgcagcagcgttcaagaagcgaagtaactgtcgtgcgc 2373
 |||||||
 Db 76562 TTCTTGATGCGCTTCTTTCACACAGGCGCAGCTGTACGAGCGCAAGTACCTGCTGCGGCC 76621
 QY 2374 aacatgatgctgt 2433
 |||||||
 Db 76622 AACATGATGCTGTGCTGTGACGACCTGTGACGAGGCGCTTGGTGTGTGTGTGTGTGTGT 76681
 QY 2434 tgcctagtatctctgtccgcgcagacactcaacagacagagcaacttcaagcctcatt 2493
 |||||||
 Db 76682 TGCTAGGTGATCTCTGCGGCCAGACCTCAACAGACAGACAGACACTTCAGGCTCCATT 76741
 QY 2494 caggaactaacgagcgcgtgcgtgcctcaactga 2526
 |||||||
 Db 76742 CAGGACTTACGAGGCGCTGCGGCTCCACCTGA 76774

RESULT 7
 AC108123/c 191481 bp DNA linear HTG 25-JAN-2002
 LOCUS AC108123/c
 DEFINITION Homo sapiens chromosome 16 clone RP11-56511, WORKING DRAFT
 SEQUENCE, 10 unordered pieces.
 ACCESSION AC108123
 VERSION AC108123.1 GI:18369969
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 191481)
TITLE DOE Joint Genome Institute.
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 191481)
TITLE DOE Joint Genome Institute.
JOURNAL Direct Submission
SUBMITTED (25-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 623518
Center clone name: RPCI-11_56511

Summary Statistics
Consensus quality: 182477 bases at least Q40
Consensus quality: 187876 bases at least Q30
Consensus quality: 188750 bases at least Q20
Estimated insert size: 192000; agarose-fp estimation
Estimated insert size: 190581; sum-of-contrigs estimation
Quality coverage: 6.71 in Q20 bases; agarose-fp estimation
Quality coverage: 6.76 in Q20 bases; sum-of-contrigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1230: contig of 1230 bp in length
1231 1330: gap of unknown length
1331 3606: contig of 2276 bp in length
3607 3706: gap of unknown length
3707 6459: contig of 2753 bp in length
6460 6559: gap of unknown length
6560 10289: contig of 3730 bp in length
10289 10389: gap of unknown length
10390 16864: contig of 6475 bp in length
16865 16964: gap of unknown length
16965 27807: contig of 10843 bp in length
27808 27907: gap of unknown length
27908 39098: contig of 11191 bp in length
39099 39198: gap of unknown length
39199 76534: contig of 37336 bp in length
76535 76634: gap of unknown length
76635 121347: contig of 44713 bp in length
121348 121447: gap of unknown length
121448 191481: contig of 70034 bp in length.
Location/Qualifiers
1. 191481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-56511"
/clone_lib="RPCI human BAC library 11"
BASE COUNT 48097 a 47075 c 46675 g 48729 t 901 others
ORIGIN
Query Match 34.9% Score 882; DB 2; Length 191481;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 932; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1594 gacctctacagatgccagcctgtgtggaagaagagtgagcactgaaggaagcagacc 1653

Db 72598 GACCTCTACAGATGCCACCTTGTGGGAAGAAGAGTGGGACCTGAGGGAAGCCAGACC 72539
QY 1654 tgcctccgcgacgtgtgtgttcttgcttctgtgtgacacactcttgggtgtctgt 1713
Db 72538 TGCCTCCCGCCACACTGT 72479
QY 1714 gcaactaacacgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1773
Db 72478 GCAGCTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 72419
QY 1774 cttagaacacccctgt 1833
Db 72418 CTAGACACCCCTGT 72359
QY 1834 ctgacagcaggtagt 1893
Db 72358 CTGGCAGCAGGTAGT 72299
QY 1894 ttgtacgcagcagcctcttgccttgccttgccttgccttgccttgccttgccttgcct 1953
Db 72298 TTGCTACGCCAGGCCCTTTTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 72239
QY 1954 cgtcatctcaactaatcatcatctcaagtcttccacaaagtactacatacttaccac 2013
Db 72238 CGCTCATTCACATATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 72179
QY 2014 gctgtgtgtcaaaacacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2073
Db 72178 GCTGT 72119
QY 2074 ctctatctcaactgt 2133
Db 72118 CTATATCTCTAATCTGT 72059
QY 2134 ttcccccactgt 2193
Db 72058 TTCCCCCATCTGT 71999
QY 2194 ttccctcaaatgt 2253
Db 71998 TTCCCTCTCAATGT 71939
QY 2254 ttgcagaggaactcaaaagaggaaggaaggaaggaaggaaggaaggaaggaagga 2313
Db 71938 TTGCCAGGAAGCTACAAAGAGGCCAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 71879
QY 2314 tccctgatacgtctctcttcaacacagcagcgtctacagagcaggtacgtctgtgtgt 2373
Db 71878 TCCGTGATCGGCTTCTTACACAGGCGCCAGGCTTACAGAGGCGAGTACTGCTGGGCC 71819
QY 2374 aacatgatgt 2433
Db 71818 AACATGATGAGCTGTGGCTGTGAGCAGCTGTGAGCAGCGGCTTGTGTGTGTGTGT 71759
QY 2434 tgcctacgt 2493
Db 71758 TGCTACGTGATCTCTGTGCGGCCAGACCTCTACAGCAGAGAGACACTTCCAGGCT 71699
QY 2494 cagagactacagagcgt 2526
Db 71698 CAGGACTACAGAGGCGGCTGTGCGGCTCCACCTGA 71666
RESULT 8
AX282898 AX282898 705 bp DNA linear PAT 02-NOV-2001
LOCUS
DEFINITION Sequence 7 from Patent WO0172842.
ACCESSION AX282898
VERSION AX282898.1 GI:16609871
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (sites)
 Walker, D.W., Scoville, J., Donoho, G., Turner, C.A., Friedrich, G.C., Abuh, A., Zambrowicz, B. and Sands, A.T.
 Novel human 74m proteins and polynucleotides encoding the same
 Patent: WO 01/2842-A 7 04-OCT-2001;
 Lexicon Genetics Incorporated (US)
 Location/Qualifiers
 source 1. 705
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 130 a 231 c 170 g 173 t 1 others
 ORIGIN

Query Match 27.9%; Score 705; DB 6; Length 705;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1822 atgctggtcctccctgagcagagtagtgcagcctctatgctctcttgggagaccaca 1881
 |||||||
 Db 1 ATGCTGGGCTCCCTGGCAGAGTAGTGAGCGCTCTATGGCTTTTGGGAAACCCACA 60
 |||||||
 QY 1882 aggcctgctgctctgctgaagcagccctcttgccttggttcaacatctctgtcc 1941
 |||||||
 Db 61 AGGCGCTGCTGCTTGTAGCGCCAGGCCCTTGTGCTTACACATCTTCTGCTCC 120
 |||||||
 QY 1942 tgcctacagctgctcattccaaactaatcatcatcttcaagtcttccacaaagtaact 2001
 |||||||
 Db 121 TGCCTACAGTTGCTGCTATTCACACTATCATATCTTCAAGTTTCCACAAAGTAACCT 180
 |||||||
 QY 2002 aacttaccacagcctgggtlccaaacacagctgctgctgttggatgatacctca 2061
 |||||||
 Db 181 ACATTTACACAGCCCTGGGGTCCAAACACAGGCTGCGCTGTTGTGATGATCACCTCA 240
 |||||||
 QY 2062 gggggccacgtcttctgcttaactgctgctgctgctgctgctgctgctgctgct 2121
 |||||||
 Db 241 GGCGCCCACTGCTTATCTTACTTAACTTGGCTGGTGTGAGCCACGCTGCTAGG 300
 |||||||
 QY 2122 gaatacagcgtctccccaatctgtgtagtctgtagtgcagagacaaactccctgggc 2181
 |||||||
 Db 301 GATACAGAGGCTTCCGCCATCTGCTGATGCTTGTAGTGCACAGACCAACTCCCTGGGC 360
 |||||||
 QY 2182 ttcatactgctcctcctctcaaatgagcctcctcctcctcctcctcctcctcctcct 2241
 |||||||
 Db 361 TTTCACTGCGCTTCTCTCAATGAGCTCTCTCCATCATGAGTCTTGTGCTGCTACCTAC 420
 |||||||
 QY 2242 ctgggtaagagacttgccagagaactacaagagagcacaatggtctacacctcagctgc 2301
 |||||||
 Db 421 CTGGGTAAAGGACTTGGCAGAGAACTACAAAGAGGCCAAATGCTACCTTACGCTGCTC 480
 |||||||
 QY 2302 ttcaactgctgctcctgagctgctcctcctcctcctcctcctcctcctcctcctcct 2361
 |||||||
 Db 481 TTCAATTGCTGCTGCTGATGCTGCTTCTTACACAGGCCAGCTCTACAGAGCAAGTAC 540
 |||||||
 QY 2362 ctgctcctgagcacaatgctgctgctgctgctgctgctgctgctgctgctgctgct 2421
 |||||||
 Db 541 CTGGCTGGGCGCAACATGATGCTGGGCTGAGCAGCCTGAGCAGCGGCTTGGTGGTAA 600
 |||||||
 QY 2422 ttctgctcctgagctgctgctcctcctcctcctcctcctcctcctcctcctcctcct 2481
 |||||||
 Db 601 TTTCTGCTCAAGTCTGATGCTGCTCTCTGCGCCAGACCTCAACAGCAGAGCACTTC 660
 |||||||
 QY 2482 caggtccatcatcagactacacagagagcgtgctgctcacttga 2526
 |||||||
 Db 661 CAGGCTCCATTTCAGACTACAGAGAGCGCTGCGCTCCACCTGA 705
 |||||||

RESULT 9
 LOCUS AC062024 185061 bp DNA linear HMG 01-SEP-2000
 DEFINITION Homo sapiens chromosome 1 clone RP11-239H20, WORKING DRAFT

ACCESSION AC062024
 VERSION AC062024.2 GI:9958200
 KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 185061)
 Waterston, R.H.
 The sequence of Homo sapiens clone
 2 (bases 1 to 185061)
 Waterston, R.H.
 Direct Submission
 Submitted (21-APR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 On Sep 1, 2000 this sequence version replaced gi:7630906.

COMMENT

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H.NH0239H20
 ----- Summary Statistics -----
 Sequencing vector: M13; 100%
 Sequencing vector: plasmid; 0%
 Chemistry: Dye-primer ET; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 168551 bases at least Q40
 Consensus quality: 174217 bases at least Q30
 Consensus quality: 177307 bases at least Q20
 Insert size: 173000; agarose-fp
 Insert size: 182161; sum-of-ctdigs
 Quality coverage: 3.82 in Q20 bases; sum-of-ctdigs
 Quality coverage: 3.68 in Q20 bases; sum-of-ctdigs
 ----- NOTE: This is a 'working draft' sequence. It currently consists of 30 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. -----
 * 1 1990: contig of 1990 bp in length
 * 1991 2090: gap of unknown length
 * 2091 3876: contig of 1786 bp in length
 * 3877 3976: gap of unknown length
 * 3977 5964: contig of 1988 bp in length
 * 5965 6064: gap of unknown length
 * 6065 7612: contig of 1548 bp in length
 * 7613 7712: gap of unknown length
 * 7712 11594: contig of 3882 bp in length
 * 11595 11694: gap of unknown length
 * 11695 15489: contig of 3795 bp in length
 * 15490 15589: gap of unknown length
 * 15590 19085: contig of 3466 bp in length
 * 19086 19185: gap of unknown length
 * 19186 21398: contig of 2213 bp in length
 * 21399 21498: gap of unknown length
 * 21499 33593: contig of 12095 bp in length
 * 33594 33693: gap of unknown length
 * 33694 36499: contig of 2806 bp in length
 * 36500 36599: gap of unknown length
 * 36600 40210: contig of 3611 bp in length
 * 40211 40310: gap of unknown length
 * 40311 43308: contig of 2998 bp in length
 * 43309 43408: gap of unknown length
 * 43409 46757: contig of 3349 bp in length

* 46758 46857: gap of unknown length
* 46858 51135: contig of 4278 bp in length
* 51136 51235: gap of unknown length
* 51236 55293: contig of 4058 bp in length
* 55294 55393: gap of unknown length
* 55394 62010: contig of 6617 bp in length
* 62011 62110: gap of unknown length
* 62111 68999: contig of 6889 bp in length
* 69000 69099: gap of unknown length
* 69100 74515: contig of 5416 bp in length
* 74516 74615: gap of unknown length
* 74616 80987: contig of 6372 bp in length
* 80988 81087: gap of unknown length
* 81088 87983: contig of 6896 bp in length
* 87984 88083: gap of unknown length
* 88084 98428: contig of 10345 bp in length
* 98429 98528: gap of unknown length
* 98529 107911: contig of 9383 bp in length
* 107912 108012: gap of unknown length
* 108012 117076: contig of 9065 bp in length
* 117077 117176: gap of unknown length
* 117177 128937: contig of 11761 bp in length
* 128938 129037: gap of unknown length
* 129038 138066: contig of 9029 bp in length
* 138067 138166: gap of unknown length
* 138167 149334: contig of 11168 bp in length
* 149335 149434: gap of unknown length
* 149435 165184: contig of 15750 bp in length
* 165185 165284: gap of unknown length
* 165285 181704: contig of 16420 bp in length
* 181705 181804: gap of unknown length
* 181805 183334: contig of 1530 bp in length
* 183335 183434: gap of unknown length
* 183435 185061: contig of 1627 bp in length.

FEATURES
source

misc_feature
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-239H20"
1..1990
/note="assembly_name:Contig11"
misc_feature
2091..3876
/note="assembly_name:Contig12"
misc_feature
3977..5964
/note="assembly_name:Contig13"
misc_feature
6065..7612
/note="assembly_name:Contig14"
misc_feature
7713..11594
/note="assembly_name:Contig15"
misc_feature
11695..15489
/note="assembly_name:Contig16"
misc_feature
15590..19085
/note="assembly_name:Contig17"
misc_feature
19186..21398
/note="assembly_name:Contig18
clone_end:SP6
vector_side:left"
misc_feature
21499..33593
/note="assembly_name:Contig36"
misc_feature
33694..36499
/note="assembly_name:Contig19"
misc_feature
36600..40210
/note="assembly_name:Contig20"
misc_feature
40311..43308
/note="assembly_name:Contig21"
misc_feature
43409..46757
/note="assembly_name:Contig22"
misc_feature
46858..51135
/note="assembly_name:Contig23"
misc_feature
51236..55293
/note="assembly_name:Contig24"
misc_feature
55394..62010

/note="assembly_name:Contig25"
misc_feature
62111..68999
/note="assembly_name:Contig26"
misc_feature
69100..74515
/note="assembly_name:Contig27"
misc_feature
74616..80987
/note="assembly_name:Contig28"
misc_feature
81088..87983
/note="assembly_name:Contig29"
misc_feature
88084..98428
/note="assembly_name:Contig30"
misc_feature
98529..107911
/note="assembly_name:Contig31"
misc_feature
108012..117076
/note="assembly_name:Contig32"
misc_feature
117177..128937
/note="assembly_name:Contig33"
misc_feature
129038..138066
/note="assembly_name:Contig34"
misc_feature
138167..149334
/note="assembly_name:Contig35"
misc_feature
149435..165184
/note="assembly_name:Contig37"
misc_feature
165285..181704
/note="assembly_name:Contig38"
misc_feature
181805..183334
/note="assembly_name:Contig8"
misc_feature
183435..185061
/note="assembly_name:Contig9"
BASE COUNT 45659 a 44829 c 45758 g 45906 t 2909 others
ORIGIN

Query Match 21.7%, Score 549, DB 2; Length 185061;
Best Local Similarity 99.7%, Pred. No. 5e-300;
Matches 719, Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1806 ccgcctgctcttcttctgctgctgctcctgcaacaggtgtagcagccctatgctt 1865
DB 33701 CCGCTGCTCTTTCTTAACTGCTGGCTCCCTGCGACAGTAGTGACCTCTAAGCTT 33760
QY 1866 ctltgggaaccacaagcctgcgtctgtctgaacagccctcttgccttgctt 1925
DB 33761 CTTTGGGAAACCAAGGCGCTGCTGCTGTGTACGACGCGCTTGGCTTGCTT 33820
QY 1926 caccatctctgtctgcctgcagcagctgcgtcattcacaactatcatcattcaagtt 1985
DB 33821 CACCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 33880
QY 1986 ttcccaaggttaacctatcttaccacagcctgggtccaaaacacagctgctgctgtt 2045
DB 33881 TTCCACCAAGTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 33940
QY 2046 tgtgtagtcagctcagcggccagctgtctatctgtcttaactgtgctgtgtgtgac 2105
DB 33941 TGTGATGATCAGCTCAGCGGCGCCAGCTGCTATCTGTCTACTTGTGTGTGTGAC 34000
QY 2106 cccaatgctctgtaggaataacacagccttcccccatctgtgtgtagctgtgacaga 2165
DB 34001 CCCACTGCTCTAGGAAATACACAGCTTCCCATCTGTGTGTGTGTGTGTGTGTGT 34060
QY 2166 gaccaactccctggtctcactgctcttccatcaaatgcccctccatcagtcgc 2225
DB 34061 GACCAACTCCCTGGGCTTCATCTGCTGCTCTCTACAAATGGCTCTCCATCATGTCG 34120
QY 2226 ctltgctgcagctacctggtgtgaagctgtgcagagaactacaagagccaaatgygt 2285
DB 34121 CTTTCTCTGACAGCTACCTGGGTAAGACTTGCCAGAGAATCAACAGAGGCCAAATGTGT 34180
QY 2286 caccctcagctcctctcactatctgttctcctgtgtgtgctcttccacacagcgcaaggt 2345
DB 34181 CACCTTACGCTGCTCTCAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 34240

QY 2346 ctacagcgcaagctaccctcgtcgagccacaatgatgctggagctgaagcagcctgaagcag 2405
|||||
Db 34241 CTACGACGGCAAGTACTCTGCTGCGG-CAACATGATGGCTGGGCTGAGCAGCCTGAGCAG 34299
QY 2406 cggcttcggtggtatcttcctgaagtcgtacgtacgtccttcgcccagacctca 2465
|||||
Db 34300 CGGCTTCGGTGGTATTTCTGCTTAAGTACGTGATCTCTGCGCCCAACCTCA 34359
QY 2466 cagcagagacacttcacagcctccatcaatcaagactaacagagcgctcgccctccacctg 2525
|||||
Db 34360 CAGCAGACAGAGCACTTCACAGGCGCTCCATTCAGAGCTACAGAGCGGCTGGGCGCTCACCTG 34419
QY 2526 a 2526
Db 34420 A 34420

RESULT 10
LOCUS G09859 270 bp DNA linear STS 15-AUG-1995
DEFINITION human STS CHLC.GCT15G02.P16729 clone GCT15G02, sequence tagged site.
ACCESSION G09859
VERSION G09859.1 GI:9411708
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human vector-pUC19 host-E.coli dut+ung+ (DH10B) Marker Selected genomic DNA prepared from xy individual of French nationality.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 270)
Murray,J., Sheffield,V., Weber,J.L., Duyk,G. and Buetow,K.H.
Cooperative Human Linkage Center
Unpublished
Synonyms: GCT15G02, CHLC.GCT15G02.T16713
Contact: Dr. Jeffrey C. Murray
UofI
The University of Iowa
Department of Pediatrics, Iowa City, IA 52242, USA
Tel: (319) 356-3508
Fax: (319) 356-3347
Email: jeff-murray@uiowa.edu
Primer A: CTCACACACAGGGGTGTCTAG
Primer B: TGGTGTTTTGGCTTTC
STS size: 124
PCR Profile:
denature: 30 seconds at 94 degrees C
annealing: 75 seconds at 55 degrees C
extension: 15 seconds at 72 degrees C
PCR cycles: 27
extension: 6 minutes at 72 degrees C
Protocol:
Template: 30ng genomic DNA
Primer: each 1.5 pmole
dNTPs: each 200 uM
Tag Polymerase: 0.3 units
Total Vol: 10 ul
Buffer:
MgCl2: 1.5mM
KCl: 50mM
Tris: 10mM
pH: 8.3
Location/Qualifiers
1..270
/organism="Homo sapiens"
/db_xref="taxon:9606"
STS
primer_bind 95..114
primer_bind complement(201..218)
BASE COUNT 66 a 90 c 75 g 39 t
ORIGIN

Query Match 10.7%; Score 270; DB 11; Length 270;
Best Local Similarity 100.0%; Pred. No. 3.3e-141;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1618 gggaaagaagatgggacacttgagggaagcagaactgtctccgcgacatgtgtgttt 1677
|||||
Db 270 GGGAAAGAAGAGTGGGACCTGTAGGAGAGCCAGACCTGCTCCCGGCACTGTGTGT 211
QY 1678 ttggtcttcgtgagacacacctttgtgtcgtcgtgagactaaagcgtgtgtgtgt 1737
|||||
Db 210 TTGGCTTGGTGAGCAGACACCTCTGGGTCTCTGGACCTAACGCTGCTGCTGCTG 151
QY 1738 ctgctcttgagactcgtcgtcgttcttgcttgacactaagacccctgttggtgagta 1797
|||||
Db 150 CTGCTCTTGGGACCTGCTGCGCTGTGCTGCTGACCTAGACACCCCTGTGTGAGGTCA 91
QY 1798 gcagggggccgcctgtgtcttcttctatgtcgtggtcccttgagcagcaggtgtgacgctc 1857
|||||
Db 90 GCAGGGGGCGCGCTGCTGCTTCTTATGCTGCGCTCCCTGGCAGCAGTGTGACACCTTC 31
QY 1858 tatgcttctttggggaaccacaagcct 1887
Db 30 TATGCTTCTTGGGAGACCCACAGGCT 1

RESULT 11
LOCUS HSHFR3501 3106 bp DNA linear PRI 27-SEP-1996
DEFINITION Human Kruppel-related 3 (HKR3) gene, exons 1-3.
ACCESSION U45324
VERSION U45324.1 GI:1177226
KEYWORDS
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3106)
Murray,J., Jensen,S.J., Sulman,E.P., Bellinger,C.P., Gates,K., Allen,C., Biegel,J.A., Brodeur,G.M. and White,P.S.
Cloning, chromosomal localization, physical mapping, and genomic characterization of HKR3
Genomics 35 (2), 289-298 (1996)
JOURNAL 96299777
MEDLINE 2 (bases 1 to 3106)
AUTHORS Marris,J.M., Jensen,S.J., Sulman,E.P., Bellinger,C.P., Gates,K., Allen,C., Biegel,J.A., Brodeur,G.M. and White,P.S.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-1996) John M. Marris, Oncology/Pediatrics, The Children's Hospital of Philadelphia, ARC 902D; 324 South 34th Street, Philadelphia, PA 19104-4318, USA
Location/Qualifiers
1..3106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
/clone="P1-4184, P1-4185, P1-4186"
685..770
/gene="HKR3"
/number=1
771..1172
/gene="HKR3"
/number=1
1173..1931
/gene="HKR3"
/number=2
1932..2746
/gene="HKR3"
/number=2
1992..1998
TATA_signal


```

      exon
      /gene="HKR3"
      2747..2988
      /gene="HKR3"
      /number=3
      Intron
      2989..>3106
      /gene="HKR3"
      /note="5' portion of Intron 3 (estimated size Intron 3 =
      3.8 kb)"
      /number=3
BASE COUNT      609 a      881 c      957 g      659 t
ORIGIN
Query Match      7.1%: Score 180; DB 9; Length 3106;
Best Local Similarity 100.0%; Pred. No. 4.3e-90;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2347 tacagcagcaagtaacctgcctgcgcgcaacatagatgctggctgagcctgagcagc 2406
      |||||||
Db 39 TAGCAGCGCAAGTACCTGCGCTGCGGCCACATGATGGCTGGGCTGAGCAGCCTGAGCAGC 98
QY 2407 ggctcgtgtggtattcttcgctcaagtgtcagtgatcctctgcgcgcagacctcaac 2466
      |||||||
Db 99 GGCTTCGGTGGGTATTTCTGCTCAAGTGTCTGATCTGCTGCGGCCAGACCTCAAC 158
QY 2467 agcacagacacactccagcgcctcatcagactacacagcagcgtgcgtgcctcacctga 2526
      |||||||
Db 159 AGCACAGACACTTCACGAGCCTCCATTCAGACTACACAGGCGCTGCGCTCCACTGA 218

RESULT 12      185061 bp      DNA      linear      HTG 01-SEP-2000
AC062024/c      Homo sapiens chromosome 1 clone Rpl1-239H20, WORKING DRAFT
LOCUS
DEFINITION
AC062024      AC062024.2      GI:9958200
AC062024      AC062024.2      GI:9958200
VERSION
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 185061)
TITLE
The sequence of Homo sapiens clone
JOURNAL
Unpublished
2 (bases 1 to 185061)
AUTHORS
Waterson,R.H.
REFERENCE
Direct Submission
Submitted (21-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On Sep 1, 2000 this sequence version replaced gi:7630906.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0239H20
----- Summary Statistics -----
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168551 bases at least Q40
Consensus quality: 174217 bases at least Q30
Consensus quality: 177307 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 182161; sum-of-contigs
Quality coverage: 3.82 in Q20 bases; agarose-fp
Quality coverage: 3.68 in Q20 bases; sum-of-contigs
```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 30 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1990: contig of 1990 bp in length
1 1991: gap of unknown length
1 2090: gap of unknown length
1 2091: contig of 1786 bp in length
1 3876: contig of 1786 bp in length
1 3877: gap of unknown length
1 3977: contig of 1988 bp in length
1 5964: gap of unknown length
1 5965: gap of unknown length
1 6064: gap of unknown length
1 6065: contig of 1548 bp in length
1 7613: gap of unknown length
1 7712: gap of unknown length
1 11594: contig of 3882 bp in length
1 11595: gap of unknown length
1 11694: contig of 3795 bp in length
1 15480: contig of 3795 bp in length
1 15489: gap of unknown length
1 15589: gap of unknown length
1 15590: gap of unknown length
1 19085: contig of 3496 bp in length
1 19086: gap of unknown length
1 21398: contig of 2213 bp in length
1 21399: gap of unknown length
1 21498: gap of unknown length
1 33593: contig of 12095 bp in length
1 33594: gap of unknown length
1 33684: gap of unknown length
1 36500: contig of 2806 bp in length
1 36509: gap of unknown length
1 40210: gap of unknown length
1 40211: contig of 3611 bp in length
1 40310: gap of unknown length
1 43308: contig of 2998 bp in length
1 43309: gap of unknown length
1 43408: gap of unknown length
1 46757: contig of 3349 bp in length
1 46758: gap of unknown length
1 46857: gap of unknown length
1 51135: contig of 4278 bp in length
1 51136: gap of unknown length
1 51235: gap of unknown length
1 55293: contig of 4058 bp in length
1 55294: gap of unknown length
1 55393: gap of unknown length
1 62010: contig of 6617 bp in length
1 62011: gap of unknown length
1 62110: gap of unknown length
1 68999: contig of 6889 bp in length
1 69000: gap of unknown length
1 69009: gap of unknown length
1 74515: contig of 5416 bp in length
1 74516: gap of unknown length
1 74615: gap of unknown length
1 80987: contig of 6372 bp in length
1 80988: gap of unknown length
1 81087: gap of unknown length
1 81088: contig of 6896 bp in length
1 87983: gap of unknown length
1 88083: gap of unknown length
1 98428: contig of 10345 bp in length
1 98429: gap of unknown length
1 98528: gap of unknown length
1 107911: contig of 9383 bp in length
1 107912: gap of unknown length
1 108011: gap of unknown length
1 108012: contig of 9065 bp in length
1 117077: gap of unknown length
1 117078: gap of unknown length
1 117176: gap of unknown length
1 128937: contig of 11761 bp in length
1 128938: gap of unknown length
1 129037: gap of unknown length
1 129038: gap of unknown length
1 138066: contig of 9029 bp in length
1 138067: gap of unknown length
1 138166: gap of unknown length
1 138167: gap of unknown length
1 149334: contig of 11168 bp in length
1 149335: gap of unknown length
1 149336: gap of unknown length
1 149434: gap of unknown length
1 15184: contig of 15750 bp in length
1 15185: gap of unknown length
1 15285: gap of unknown length
1 165185: gap of unknown length
1 165285: gap of unknown length
1 181705: contig of 16420 bp in length
1 181706: gap of unknown length
1 181805: gap of unknown length
1 183335: contig of 1530 bp in length
1 183336: gap of unknown length
1 183435: gap of unknown length
1 183436: contig of 1627 bp in length.
location/Qualifiers
1.185061
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
```

```
misc_feature /clone="RP11-239H20"
1..1990
/note="assembly_name:Contig11"
misc_feature 2091..3876
/note="assembly_name:Contig12"
misc_feature 3977..5964
/note="assembly_name:Contig13"
misc_feature 6065..7612
/note="assembly_name:Contig14"
misc_feature 7713..11594
/note="assembly_name:Contig15"
misc_feature 11695..15489
/note="assembly_name:Contig16"
misc_feature 15590..19085
/note="assembly_name:Contig17"
misc_feature 19186..21398
/note="assembly_name:Contig18"
clone_end:SP6
vector_side:left"
21499..33593
/note="assembly_name:Contig36"
misc_feature 33694..36499
/note="assembly_name:Contig19"
misc_feature 36600..40210
/note="assembly_name:Contig20"
misc_feature 40311..43308
/note="assembly_name:Contig21"
misc_feature 43409..46757
/note="assembly_name:Contig22"
misc_feature 46858..51135
/note="assembly_name:Contig23"
misc_feature 51236..55293
/note="assembly_name:Contig24"
misc_feature 55394..62010
/note="assembly_name:Contig25"
misc_feature 62111..68999
/note="assembly_name:Contig26"
misc_feature 69100..74515
/note="assembly_name:Contig27"
misc_feature 74616..80987
/note="assembly_name:Contig28"
misc_feature 81088..87983
/note="assembly_name:Contig29"
misc_feature 88084..98428
/note="assembly_name:Contig30"
misc_feature 98529..107911
/note="assembly_name:Contig31"
misc_feature 108012..117076
/note="assembly_name:Contig32"
misc_feature 117177..128937
/note="assembly_name:Contig33"
misc_feature 129038..138066
/note="assembly_name:Contig34"
misc_feature 138167..149334
/note="assembly_name:Contig35"
misc_feature 149435..165184
/note="assembly_name:Contig37"
misc_feature 165285..181704
/note="assembly_name:Contig38"
misc_feature 181805..183334
/note="assembly_name:Contig8"
misc_feature 183435..185061
/note="assembly_name:Contig9"
BASE COUNT 45659 a 44829 c 45758 g 45906 t 2909 others
ORIGIN
```

```
Query Match 6.8%; Score 172; DB 2; Length 185061;
Best Local Similarity 100.0%; Pred. No. 1.2e-85;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 22 ctgtgtgcctcagcttctcattctctctctgtggtctgtgctgcacatgacagag 81
|||||
```

```
Db 37929 CTGTGCGGCGCTCAGCTTCATTTCTGCTGCTG66GCTTTGCCATACACGAG 37870
QY 82 tcttctcctgactaacctcccgagagattactctctgagcgtgtctctccat 141
|||||
Db 37869 TCTTCTCCGATTCACTCCCTCCCGAGATTTACTCTCGCAGGCGCTGTCTCCAT 37810
|||||
QY 142 tctgtctgtctgcaggtgagcagacagaccgagtgacctgtgtacaggt 193
|||||
Db 37809 TCTGGCTGTCTCAGGTGAGCAGACAGACCGAGGTGACCTGTGACAGGT 37758
|||||

RESULT 13
AC108123 191481 bp DNA linear HTG 25-JAN-2002
AC108123 Homo sapiens chromosome 16 clone RP11-56511, WORKING DRAFT
LOCUS SEQUENCE, 10 unordered pieces.
ACCESSION AC108123.1 GI:18369969
VERSION AC108123.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 191481)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 191481)
DOE Joint Genome Institute.
Direct Submission
Submitted (25-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 623518
Center clone name: RPC1-11_56511
-----
Summary Statistics
Consensus quality: 182477 bases at least Q40
Consensus quality: 187876 bases at least Q30
Consensus quality: 188750 bases at least Q20
Estimated insert size: 192000; agarose-fp estimation
Estimated insert size: 190581; sum-of-contigs estimation
Quality coverage: 6.71 in Q20 bases; agarose-fp estimation
Quality coverage: 6.76 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1230: contig of 1230 bp in length
* 1231 1330: gap of unknown length
* 1331 3606: contig of 2276 bp in length
* 3607 3706: gap of unknown length
* 3707 6459: contig of 2753 bp in length
* 6460 6559: gap of unknown length
* 6560 10289: contig of 3730 bp in length
* 10290 10389: gap of unknown length
* 10390 16864: contig of 6475 bp in length
* 16865 16965: gap of unknown length
* 16965 27807: contig of 10843 bp in length
* 27808 27907: gap of unknown length
* 27908 39098: contig of 11191 bp in length
* 39099 39199: gap of unknown length
* 39199 76534: contig of 37336 bp in length
* 76535 76634: gap of unknown length
```


PE 28-MAR-2001; 2001MO-US09996.
 PR 28-MAR-2000; 2000US-192978P.
 XX (LEXI-) LEXICON GENETICS INC.
 PA
 PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
 PI zambrowicz B, Sands AT;
 XX
 DR WPI: 2001-616474/71.
 XX P-PSDB; AAE11969.
 PT Novel isolated polynucleotides encoding human G-protein coupled
 PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
 PS
 PS Claim 1; Page 71-72; 80pp; English.
 XX
 XX The present sequence is a cDNA encoding human novel G-protein coupled
 CC receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
 CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeats rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.
 XX
 XX Sequence 2526 BP; 469 A; 770 C; 697 G; 587 T; 3 other;

Query Match	100.0%	Score 2526	DB 22	Length 2526
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2526	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0
QY 1	atgcgtcctcgtacagcgctcgcctggtctgcgcgcgcagctcttcattcttcgtcgtcgtgacc	60		
Db 1	atgcgtcctcgtacagcgctcgcctggtctgcgcgcgcagctcttcattcttcgtcgtcgtgacc	60		
QY 61	tttgccctgcataagaagaaggagctctctctcgtaccttcaacctctccgcggagattacctctg	120		
Db 61	tttgccctgcataagaagaaggagctctctctcgtaccttcaacctctccgcggagattacctctg	120		
QY 121	gcagcgccgttcctccctccatctcgtgcgtctcgcaggtgtaggacagacggcgaagtgtacc	180		
Db 121	gcagcgccgttcctccctccatctcgtgcgtctcgcaggtgtaggacagacggcgaagtgtacc	180		
QY 181	ctgtgtgcacaggtctctgtacgttcaatgaagaatggtctaacacctctctccaggtatgcgg	240		
Db 181	ctgtgtgcacaggtctctgtacgttcaatgaagaatggtctaacacctctctccaggtatgcgg	240		
QY 241	cttgagggtttgaagaataaacaacaccgcacgcgcctcgtgcccacaataacccttgggtacc	300		
Db 241	cttgagggtttgaagaataaacaacaccgcacgcgcctcgtgcccacaataacccttgggtacc	300		
QY 301	cagctgtatgatgtgtcttctgtacacctcgcgaatgtgtatgacacagctagagatgtctcc	360		
Db 301	cagctgtatgatgtgtcttctgtacacctcgcgaatgtgtatgacacagctagagatgtctcc	360		
QY 361	ctgccaagggaacaacacacatagagctccaagaagaccttccactatctccctcaggtg	420		
Db 361	ctgccaagggaacaacacacatagagctccaagaagaccttccactatctccctcaggtg	420		
QY 421	ctggcagtgatgtggcctcgcacagacacaaacctgtgccaacacacagccgacctgtgagc	480		
Db 421	ctggcagtgatgtggcctcgcacagacacaaacctgtgccaacacacagccgacctgtgagc	480		

[illegible]

```
Dh 1501 tgtcttgagggaaccagcgagtggtacgggttcacatcactgctgttgatgtgtg 1560
Qy 1561 cccgtgtgggttgagaccttctcaacaagtgacctctacagatgcacacctgtggg 1620
Db 1561 cccgtgtgggttgagaccttctcaacaagtgacctctacagatgcacacctgtggg 1620
Qy 1621 aaagaaggttgagcccggaagcagacacgtctcccgacgtgtgtgtttg 1680
Db 1621 aaagaaggttgagcccggaagcagacacgtctcccgacgtgtgtgtttg 1680
Qy 1681 gcttgcgtgagcaacacctcttgggtgctgtgcaagtaacaagctgtctgtctgtg 1740
Db 1681 gcttgcgtgagcaacacctcttgggtgctgtgcaagtaacaagctgtctgtctgtg 1740
Qy 1741 ctgcttgggaactgtgctgtctgtgctgtgcaactagacacctgtgtgtagtgacga 1800
Db 1741 ctgcttgggaactgtgctgtctgtgctgtgcaactagacacctgtgtgtagtgacga 1800
Qy 1801 gggggccgctgtcttctctatgtggtggtcccttgagcaggttagtgagcctctat 1860
Db 1801 gggggccgctgtcttctctatgtggtggtcccttgagcaggttagtgagcctctat 1860
Qy 1861 ggtctcttggggaaccacaaagcctgctgtgtgtacgacagccctcttggccct 1920
Db 1861 ggtctcttggggaaccacaaagcctgctgtgtgtacgacagccctcttggccct 1920
Qy 1921 ggttcaacatcttcctgctgtctgtgctgaagtgtgctatctcaactatcatcttc 1980
Db 1921 ggttcaacatcttcctgctgtctgtgctgaagtgtgctatctcaactatcatcttc 1980
Qy 1981 aagtttccacaaagtgtaacatcttcaacacgctgtggtccaaacacagtgctgac 2040
Db 1981 aagtttccacaaagtgtaacatcttcaacacgctgtggtccaaacacagtgctgac 2040
Qy 2041 ctgttctgtatgatcaagctcagcgccagctgtctatctgtctcaacttggctgtgtg 2100
Db 2041 ctgttctgtatgatcaagctcagcgccagctgtctatctgtctcaacttggctgtgtg 2100
Qy 2101 tggagcccaactgctgtcaggaataacacagcgttcccccctctgtgtgtgtgtgtg 2160
Db 2101 tggagcccaactgctgtcaggaataacacagcgttcccccctctgtgtgtgtgtgtg 2160
Qy 2161 acaagaacaaactccctggtgtcatactggtcctctcaaatggtcctctccatc 2220
Db 2161 acaagaacaaactccctggtgtcatactggtcctctcaaatggtcctctccatc 2220
Qy 2221 agtgccttctgctcagctacacgtggtgaagacttgcagagaactacaacagagccaa 2280
Db 2221 agtgccttctgctcagctacacgtggtgaagacttgcagagaactacaacagagccaa 2280
Qy 2281 tgytcaacctcagctgtgtcttcaactgtgtctgtgctgtgctcttcaacagagcc 2340
Db 2281 tgytcaacctcagctgtgtcttcaactgtgtctgtgctgtgctcttcaacagagcc 2340
Qy 2341 agcgtctacagcagcaagtaactcgtcgtgcgcacaaagatggtgcgtgagcagctg 2400
Db 2341 agcgtctacagcagcaagtaactcgtcgtgcgcacaaagatggtgcgtgagcagctg 2400
Qy 2401 agcagcggtcctcggtgggtatcttctgcctaaagtacgtacgtatcctctgcgcagac 2460
Db 2401 agcagcggtcctcggtgggtatcttctgcctaaagtacgtacgtatcctctgcgcagac 2460
Qy 2461 ctcaacaacagagcaacttccagcctccatcagagactacaagagcgctgagcctcc 2520
Db 2461 ctcaacaacagagcaacttccagcctccatcagagactacaagagcgctgagcctcc 2520
Qy 2521 acctga 2526
Db 2521 acctga 2526
```

RESULT 2
AAD19505

```
ID AAD19505 standard; DNA: 2951 BP.
XX
AC AAD19505;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human novel G-protein coupled receptor DNA with 5' and 3' regions.
XX
KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW antisense-therapy; signal transduction; behavioural disorder; obesity;
KW heartbeat rate; inflammation; immune disorder; diabetes; cancer;
KW coronary disease; ds.
XX
OS Homo sapiens.
XX
PN WO200172842-A2.
XX
PD 04-OCT-2001.
XX
PE 28-MAR-2001; 2001WO-US09996.
XX
PR 28-MAR-2000; 2000US-192978P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
PI WPI: 2001-616474/71.
XX
PT Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
XX
PS Disclosure: Page 79-80; 80pp; English.
XX
CC The present sequence is human novel G-protein coupled receptor (NGPCR)
CC DNA with 5' and 3' regions. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeat rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.
XX
SQ Sequence 2951 BP; 559 A; 888 C; 831 G; 669 T; 4 other:
XX
```

Query Match 100.0%; Score 2526; DB 22; Length 2951;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 atgcgtcctgacagcgcgtcgtgtgctgagcgttctcatcttctgtcgtgagcc 60
Db 232 atgcgtcctgacagcgcgtcgtgtgctgagcgttctcatcttctgtcgtgagcc 291
Qy 61 ttgcctgcataagcagagagctctctcactcacccctcccgagattactctctg 120
Db 292 ttgcctgcataagcagagagctctctcactcacccctcccgagattactctctg 351
Qy 121 gaagcctgttccctctccatctgtcgtcgtcaggtgagagcacagaccgaggtgacc 180
Db 352 gaagcctgttccctctccatctgtcgtcgtcaggtgagagcacagaccgaggtgacc 411
```

QY 181 ctgtgtgacaggtctgttagctcaatgacatggtctaccacctcttcaggtatgctg 240
|||||
Db 412 cgtgtgacaggtctgttagctcaatgacatggtctaccacctcttcaggtatgctg 471
QY 241 ctgtgtgttgagagataaacaactccacggtctgtgtcccaactaacctgggttac 300
|||||
Db 472 ctgtgtgttgagagataaacaactccacggtctgtgtcccaactaacctgggttac 531
QY 301 cagctgtatgattgtgtgtgtgactctgtccaatgtgtatgtccacgtctgagtgctcc 360
|||||
Db 532 cagctgtatgattgtgtgtgtgactctgtccaatgtgtatgtccacgtctgagtgctcc 591
QY 361 ctgtccagggacaacacacatagagctccaaaggagacctctccactattccctacgtgt 420
|||||
Db 592 ctgtccagggacaacacacatagagctccaaaggagacctctccactattccctacgtgt 651
QY 421 ctgtccaggtattgtgtgtgtgactccacgaacctgtgtccacacacacggtctgtgagc 480
|||||
Db 652 ctgtccaggtattgtgtgtgtgactccacgaacctgtgtccacacacacggtctgtgagc 711
QY 481 ccttctctgtgtcccaatgattagctatgtcgtccagcagcgagacgtctcagcgtgagcg 540
|||||
Db 712 ccttctctgtgtcccaatgattagctatgtcgtccagcagcgagacgtctcagcgtgagcg 771
QY 541 cagatccctcttctctgtccac 600
|||||
Db 772 cagatccctcttctctgtccac 831
QY 601 ctgtgtgtgtcagaagtctgt 660
|||||
Db 832 ctgtgtgtgtcagaagtctgt 891
QY 661 gggcagctaggggtgtgagcactgtgagacacacacacacacacacacacacacacacac 720
|||||
Db 892 gggcagctaggggtgtgagcactgtgagacacacacacacacacacacacacacacacac 951
QY 721 ttcacagacatcatgtccctctctgtccacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
|||||
Db 952 ttcacagacatcatgtccctctctgtccacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1011
QY 781 cgtccacgtgtccacagcgt 840
|||||
Db 1012 cgtccacgtgtccacagcgt 1071
QY 841 aggtgttcttctcagatcgt 900
|||||
Db 1072 aggtgttcttctcagatcgt 1131
QY 901 gaagcctgtgtccctctccagacacatcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
|||||
Db 1132 gaagcctgtgtccctctccagacacatcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1191
QY 961 ggt 1020
|||||
Db 1192 ggt 1251
QY 1021 tctgtccgtgtcagacagagagcccttaggtctgtccacaaagggtctcgtgtgtgtgtgt 1080
|||||
Db 1252 tctgtccgtgtcagacagagagcccttaggtctgtccacaaagggtctcgtgtgtgtgtgt 1311
QY 1081 aatcagctctgtcagagatgt 1140
|||||
Db 1312 aatcagctctgtcagagatgt 1371
QY 1141 tctctcagatgt 1200
|||||
Db 1372 tctctcagatgt 1431
QY 1201 caacagctctgt 1260
|||||
Db 1432 caacagctctgt 1491

QY 1261 ctttggagcagatcccaacagtgatcttctctacacaaagacgtgtgtgtttaat 1320
|||||
Db 1492 ctttggagcagatcccaacagtgatcttctctacacaaagacgtgtgtgtttaat 1551
QY 1321 gacaacagagatccctctcagtgatctaaacatattgtgtgtgtgtgtgtgtgtgtgtgtgt 1380
|||||
Db 1552 gacaacagagatccctctcagtgatctaaacatattgtgtgtgtgtgtgtgtgtgtgtgtgt 1611
QY 1381 tggacctcaaggtctctgt 1440
|||||
Db 1612 tggacctcaaggtctctgt 1671
QY 1441 accaaatccagtgatgt 1500
|||||
Db 1672 accaaatccagtgatgt 1731
QY 1501 tgtcttgaaagggtcagacagagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560
|||||
Db 1732 tgtcttgaaagggtcagacagagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1791
QY 1561 cctgt 1620
|||||
Db 1792 cctgt 1851
QY 1621 aaagaagatgt 1680
|||||
Db 1852 aaagaagatgt 1911
QY 1681 gcttgcgt 1740
|||||
Db 1912 gcttgcgt 1971
QY 1741 ctgt 1800
|||||
Db 1972 ctgt 2031
QY 1801 ggggt 1860
|||||
Db 2032 ggggt 2091
QY 1861 ggttcttctgt 1920
|||||
Db 2092 ggttcttctgt 2151
QY 1921 ggttccacatcttctgt 1980
|||||
Db 2152 ggttccacatcttctgt 2211
QY 1981 aagtttccacaaagttacatcatctacacacacacacacacacacacacacacacacacac 2040
|||||
Db 2212 aagtttccacaaagttacatcatctacacacacacacacacacacacacacacacacacac 2271
QY 2041 ctgt 2100
|||||
Db 2272 ctgt 2331
QY 2101 tggacccac 2160
|||||
Db 2332 tggacccac 2391
QY 2161 aacagac 2220
|||||
Db 2392 aacagac 2451
QY 2221 agtgccttgcctgtcagac 2280
|||||
Db 2452 agtgccttgcctgtcagac 2511
QY 2281 tgytcaacctcaagcctgtcttcaacctcgtgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2340
|||||
Db 2512 tgytcaacctcaagcctgtcttcaacctcgtgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2571
QY 2341 agcgttcaagcaggaatcactgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2400


```

Db 761 gaagcctatgcccgggacagagaaggccctagagccttgccacaaggctcctggtgc 840
Oy 1075 agcaacaatcaagctctgcagagaatgcacaagcttcatatgcacacagatgccaaagctc 1134
Db 841 agcagcaatcagctctgcagagaatgcacaagcttcatatgcacacagatgccaaagctc 900
Oy 1135 aaagcctctccatgagtctgctctacaacagcataccgggctggtatgctggcccat 1194
Db 901 aaagcctctccatgagtctgctctacaacagcataccgggctggtatgctggcccat 960
Oy 1195 ggcctccaaacagctcctggtgctgctctgagcttcttccaggggccggaattacccc 1254
Db 961 ggcctccaaacagctcctggtgctgctctgagcttcttccaggggccggaattacccc 1020
Oy 1255 tggcagcttttgagcaagatccacaagatgcatcttcccttcaacaagaagacctgtggcg 1314
Db 1021 tggcagcttttgagcaagatccacaagatgcatcttcccttcaacaagaagacctgtggcg 1080
Oy 1315 tttaatgacaacagagatccctctcagtagctataacataatgtcctgagactggaatgga 1374
Db 1081 tttaatgacaacagagatccctctcagtagctataacataatgtcctgagactggaatgga 1140
Oy 1375 cccaagtgcactctcaagctctcctgcttccctccacatggtctccaggttcaagtaaacata 1434
Db 1141 cccaagtgcactctcaagctctcctgcttccctccacatggtctccaggttcaagtaaacata 1200
Oy 1435 aatgagacaaatccagatgagcagcgaagaaagacaacacaggtgctcctaagtctgtgttc 1494
Db 1201 aatgagacaaatccagatgagcagcgaagaaagacaacacaggtgctcctaagtctgtgttc 1260
Oy 1495 agcagatgctctgaaaggagcagcagcagatggttaacgggtttccatcaactgctgtctgag 1554
Db 1261 agcagatgctctgaaaggagcagcagcagatggttaacgggtttccatcaactgctgtctgag 1320
Oy 1555 tgtgtgcccctgtggggctgagagccttccccaagaagatgactctacagatgagcagcct 1614
Db 1321 tgtgtgcccctgtggggctgagagccttccccaagaagatgactctacagatgagcagcct 1380
Oy 1615 tgtgtgaaagaaagatgagcagcctgaggaagcagacactgtctcccgacactgtgtg 1674
Db 1381 tgtgtgaaagaaagatgagcagcctgaggaagcagacactgtctcccgacactgtgtg 1440
Oy 1675 ttttggcttggctgagcagacacactctggtgtgctgctgagcagtaaacgctgctgtc 1734
Db 1441 ttttggcttggctgagcagacacactctggtgtgctgctgagcagtaaacgctgctgtc 1500
Oy 1735 ctgagctgctgagcagcctgctgctgttggcttggccttgcacatgcacacccctgtgtgag 1794
Db 1501 ctgagctgctgagcagcctgctgctgttggcttggccttgcacatgcacacccctgtgtgag 1560
Oy 1795 tcaagcaggggagcagcctgtgcttcttcatgctgagcctccctgagcagagtagtgcagc 1854
Db 1561 tcaagcaggggagcagcctgtgcttcttcatgctgagcctccctgagcagagtagtgcagc 1620
Oy 1855 cctcatgctcttcttggggaaccccaagcctgctgctgtctgctacgagcagcctctt 1914
Db 1621 cctcatgctcttcttggggaaccccaagcctgctgctgtctgctacgagcagcctctt 1680
Oy 1915 ggccttggcttaccacatctctgtccctgctgagcaggttcgctcatctccaataatcac 1974
Db 1681 ggccttggcttaccacatctctgtccctgctgagcaggttcgctcatctccaataatcac 1740
Oy 1975 atcttcaagtcttccacaagaagtaacctatacttcaacagcctggtgtccaaaacacag 2034
Db 1741 atcttcaagtcttccacaagaagtaacctatacttcaacagcctggtgtccaaaacacag 1800
Oy 2035 gctgagccttcttggatgctcagctcagcggccagcctgtctatctgttacttgctg 2094
Db 1801 gctgagccttcttggatgctcagctcagcggccagcctgtctatctgttacttgctg 1860
Oy 2095 gtggtgtgagcccaactgctctagagataccagcgttcccccattgtgtgagct 2154
Db 1861 gtggtgtgagcccaactgctctagagataccagcgttcccccattgtgtgagct 1920

```

```

Oy 2155 gagtgcacagagaccacactccctgggcttcatatgctctctcttacaatggctctc 2214
Db 1921 gagtgcacagagaccacactccctgggcttcatatgctctctcttacaatggctctc 1980
Oy 2215 tccatcagtgctcttggctgcagcactggtgtaaggacttgcagaaactacaacagag 2274
Db 1981 tccatcagtgctcttggctgcagcactggtgtaaggacttgcagaaactacaacagag 2040
Oy 2275 gccaatggttacccttcaagcctgtcttcaacttcgttccctgtagctccttccac 2334
Db 2041 gccaatggttacccttcaagcctgtcttcaacttcgttccctgtagctccttccac 2100
Oy 2335 acggcagcgttctacagcagcgaagtaacctgctgcggccaatgatatgctgctgagc 2394
Db 2101 acggcagcgttctacagcagcgaagtaacctgctgcggccaatgatatgctgctgagc 2160
Oy 2395 agcctgagcagcgttctggtggtatcttctgctaaagtgtgtaagtatctctgcgcgc 2454
Db 2161 agcctgagcagcgttctggtggtatcttctgctaaagtgtgtaagtatctctgcgcgc 2220
Oy 2455 ccagacctcaacagcagcagacacttccagcctccattcaagactacaacagagcgtgc 2514
Db 2221 ccagacctcaacagcagcagacacttccagcctccattcaagactacaacagagcgtgc 2280
Oy 2515 ggtccacactga 2526
Db 2281 ggtccacactga 2292

RESULT 4
AADI17516
ID AADI17516 standard; cDNA; 2526 BP.
XX
AC AADI17516;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human taste receptor, hTIR1 cDNA coding sequence.
XX
KW Human; taste-cell-specific G protein-coupled receptor; hTIR1; drug;
KW genetic modulation; pharmaceutical; taste sensation; food industry;
KW chemosensory transduction; ss.
XX
OS Homo sapiens.
XX
FH Key 1.2526 Location/Qualifiers
FT CDS
FT
FT /tag= a
FT /product= "Human taste receptor, hTIR1 protein"
FT /transl_except= (pos:820..823, aa:Phe)
FT /note= "This codon has an apparent 1 nucleotide insertion
FT which alters the reading frame"
XX
PN WO20016563-A2.
XX
PD 13-SEP-2001.
XX
PE 07-MAR-2001; 2001WO-US07265.
XX
PR 07-MAR-2000; 2000US-0187546.
PR 07-APR-2000; 2000US-0195536.
PR 06-JUN-2000; 2000US-0209840.
PR 23-JUN-2000; 2000US-0214213.
PR 17-AUG-2000; 2000US-0226448.
PR 03-JAN-2001; 2001US-0259227.
XX
PA (SEMO-) SEMOMTX INC.
XX
PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
XX
DR WPI: 2001-582267/65.
DR P-PSDB: AAE10372.

```



```

Db 1740 gctgcttgagactgctgctgttgcctgcgcacacagacacccctgtgtgagcagc 1799
Qy 1800 aaggagcgccctgtgtcttcttatgtgtgctccctgcagcagcgtatgtgcagccttca 1859
Db 1800 aaggagcgccctgtgtcttcttatgtgtgctccctgcagcagcgtatgtgcagccttca 1859
Qy 1860 tggctctcttgaggaaacacacagagcctgtgtctgtctacagcagccctcttgccct 1919
Db 1860 tggctctcttgaggaaacacacagagcctgtgtctgtctacagcagccctcttgccct 1919
Qy 1920 tggcttaccatctctcctgtcctgcctgcagagctgcctcatcctcaatcatcatctt 1979
Db 1920 tggcttaccatctctcctgtcctgcctgcagagctgcctcatcctcaatcatcatctt 1979
Qy 1980 caagtttccacacaaagtaacctacattctacacagcctgtgtgtccaaacacagctgtctg 2039
Db 1980 caagtttccacacaaagtaacctacattctacacagcctgtgtgtccaaacacagctgtctg 2039
Qy 2040 cctgttctgatatcatcagctcagcagcagcctgtctatctgtctactgtgtgtgt 2099
Db 2040 cctgttctgatatcatcagctcagcagcagcctgtctatctgtctactgtgtgtgt 2099
Qy 2100 gtggaccacactgtcctgtcaggaataacacagcgtctcccccactgtgtgtgtgt 2159
Db 2100 gtggaccacactgtcctgtcaggaataacacagcgtctcccccactgtgtgtgtgt 2159
Qy 2160 caaagagacaaactcctcctgagcttactgtcctcctcctacatgagcctcctccat 2219
Db 2160 caaagagacaaactcctcctgagcttactgtcctcctcctacatgagcctcctccat 2219
Qy 2220 cagtgctcttgcctgcagcctacacacagcctgtcagcagacacacacagcagc 2279
Db 2220 cagtgctcttgcctgcagcctacacacagcctgtcagcagacacacacagcagc 2279
Qy 2280 atggtgtacactttagcctgtccttcaacttgcctgtcctgtgtgtgtgtgtgtgtgt 2339
Db 2280 atggtgtacactttagcctgtccttcaacttgcctgtcctgtgtgtgtgtgtgtgtgt 2339
Qy 2340 cagcgtctacagcagcagacacacacacacacacacacacacacacacacacacacac 2399
Db 2340 cagcgtctacagcagcagacacacacacacacacacacacacacacacacacacacac 2399
Qy 2400 gagcagcgtctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2459
Db 2400 gagcagcgtctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2459
Qy 2460 cctcaacacagacagacacacacacacacacacacacacacacacacacacacac 2519
Db 2460 cctcaacacagacagacacacacacacacacacacacacacacacacacacacac 2519
Qy 2520 cactcga 2526
Db 2520 cactcga 2526

```

```

RESULT 5
AAZ50745
ID AAZ50745 standard; cDNA; 2333 BP.
XX
AC AAZ50745;
XX
XX 31-MAY-2000 (first entry)
DE Human sensory transduction G-protein coupled receptor-B3 cDNA.
XX
XX Human; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
XX sensory cell; taste receptor cell; screen; taste modulator;
XX pharmaceutical; food industry; taste topographic map; tongue; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers

```

```

FT CDS 1..2333
FT /tag= a
FT /product= "GPCR-B3"
FT /trans_except= (pos:865..866, aa:His)
PN WO200006592-A1.
XX
XX 10-FEB-2000.
XX
XX 27-JUL-1999; 99WO-US17099.
XX
XX 28-JUL-1998; 98US-0094465.
XX
XX (REGC ) UNIV CALIFORNIA.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
XX
XX WPI; 2000-205451/18.
XX P-PSDB; AAY45023.
XX
XX New isolated sensory transduction G-protein coupled receptor, useful
XX for developing products for use in studying and modulating the taste
XX transduction pathway
XX
XX Claim 5; Pages 79-80; 83pp; English.
XX
XX The present sequence is a cDNA encoding a taste cell specific G-protein
XX coupled receptor, GPCR-B3 which is involved in sensory transduction.
XX This sequence was isolated from the human testis library.
XX GPCR-B3 is specifically expressed in foliate and fungiform cells, with
XX lower expression in circumvallate taste receptor cells of the tongue.
XX The present sequence is used to screen compounds that modulate sensory
XX signalling in taste cells, especially taste modulators useful in
XX pharmaceutical and food industries to customise taste. The sequence
XX can also be used as probe for identifying taste cells and
XX subsets of taste receptor cells such as foliate, fungiform and
XX circumvallate. Such probes are also useful to generate taste
XX topographic maps that elucidate the relationship between the taste
XX cells of the tongue and sensory neurons leading to taste centres
XX in the brain.
XX
XX Sequence 2333 BP; 445 A; 705 C; 646 G; 537 T; 0 other;
SQ

```

```

Query Match 41.5%; Score 1048; DB 21; Length 2333;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1808; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

```

```

Qy 709 atctgattgcttcaagacatcatgcccctctcctgcagcagtggtgagatgagatg 768
Db 517 atctgattgcttcaagacatcatgcccctctcctgcagcagtggtgagatgagatg 576
Qy 769 caatgctcatatgcccacccctgcagcagcagcagcagcagcagcagcagcagcagcagc 828
Db 577 caatgctcatatgcccacccctgcagcagcagcagcagcagcagcagcagcagcagcagc 636
Qy 829 cggcagctggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 888
Db 637 cggcagctggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 696
Qy 889 tgggtcgcctcaagaagcctgtgcccctctcagcagcagcagcagcagcagcagcagcag 948
Db 697 tgggtcgcctcaagaagcctgtgcccctctcagcagcagcagcagcagcagcagcagcag 756
Qy 949 cggattggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1008
Db 757 cggattggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 816
Qy 1009 ttgtaagaagcctatcccgcgagcagcagcagcagcagcagcagcagcagcagcagcagc 1068
Db 817 ttgtaagaagcctatcccgcgagcagcagcagcagcagcagcagcagcagcagcagcagc 875

```

QY 1069 tggtagcagcaacalacagctctgacagaaatgccaagcttccatgtrcacacagatgcc 1128
|||||
Db 876 tggtagcagcaacalacagctctgacagaaatgccaagcttccatgtrcacacagatgcc 935
QY 1129 aagctcaagccttccatgagtgctgcttcaacaacgataccgggctgtgtatggctg 1188
Db 936 aagctcaagccttccatgagtgctgcttcaacaacgataccgggctgtgtatggctg 995
QY 1189 gcccttgccctccacacagctctcctgagctgtgcttccctctgagct-tgttccaaaggccgagct 1247
Db 996 gcccttgccctccacacagctctcctgagctgtgcttccct-gagctctgttccaaaggccgagct 1054
QY 1248 ctacccctgagcagcttcttgagcagatccacaaagtgatcttcccttccacaagaagac 1307
Db 1055 ctacccctgagcagcttcttgagcagatccacaaagtgatcttcccttccacaagaagac 1114
QY 1308 tgtgagctttaaatacaacaagatccctcagtaagctataataatgtccctggagctg 1367
Db 1115 tgtgagctttaaatacaacaagatccctcagtaagctataataatgtccctggagctg 1174
QY 1368 gaatggaccacaaatgagccttccagctctcgtgttccctccacatggttccagttcaact 1427
Db 1175 gaatggaccacaaatgagccttccagctctcgtgttccctccacatggttccagttcaact 1234
QY 1428 aaacataaatgagacccaacacagtgagcagaaagagacacacagtgctcaagtctgt 1487
Db 1235 aaacataaatgagacccaacacagtgagcagaaagagacacacagtgctcaagtctgt 1294
QY 1488 gtgttccagcagctgtcttgaaaggacacagcagtggttaacgggttccatcaactgtg 1547
Db 1295 gtgttccagcagctgtcttgaaaggacacagcagtggttaacgggttccatcaactgtg 1354
QY 1548 ctttgagtggttgccctgttgagtggttgagccttccaaagagtgagccttccaaagt 1607
Db 1355 ctttgagtggttgccctgttgagtggttgagccttccaaagagtgagccttccaaagt 1414
QY 1608 ccagccttgttggaagaagagtgagcagctgaggaagagacacacttcccccgagac 1667
Db 1415 ccagccttgttggaagaagagtgagcagctgaggaagagacacacttcccccgagac 1474
QY 1668 tgtggtgttcttgagcttgcttgagcacactcttggtgtgctgctgagcagctaaacgt 1727
Db 1475 tgtggtgttcttgagcttgcttgagcacactcttggtgtgctgctgagcagctaaacgt 1534
QY 1728 gcttctgctgtgctgtgcttgagcagctgtgctgttcttgccggacacactaagacacccctgt 1787
Db 1535 gcttctgctgtgctgtgcttgagcagctgtgctgttcttgccggacacacccctgt 1594
QY 1788 ggtgagtgcaagagggcgccgctgtgcttcttatagtctgggctccctggcagcagtag 1847
Db 1595 ggtgagtgcaagagggcgccgctgtgcttcttatagtctgggctccctggcagcagtag 1654
QY 1848 tggcagcctctatgcttcttggggaacccacaagcctgtgctgtgtaagcagcagc 1907
Db 1655 tggcagcctctatgcttcttggggaacccacaagcctgtgctgtgtaagcagcagc 1714
QY 1908 ccttcttgcccttggttccacacatcttcttctgctgctgagcagtgcttccatccact 1967
Db 1715 ccttcttgcccttggttccacacatcttcttctgctgctgagcagtgcttccatccact 1774
QY 1968 aatcatcatcttcaagtttccacaagaagtacatacttccacaagcgtgtgtccaaa 2027
Db 1775 aatcatcatcttcaagtttccacaagaagtacatacttccacaagcgtgtgtgtccaaa 1834
QY 2028 ccacggtagctgagctgttggatgatacagctcagcgccagcagctgtatctgtctaac 2087
Db 1835 ccacggtagctgagctgttggatgatacagctcagcgccagcagctgtatctgtctaac 1894
QY 2088 ttggctggtggttgagacccaacgctgtcaggaagaatacagagccttccccacatctgt 2147
Db 1895 ttggctggtggttgagacccaacgctgtcaggaagaatacagagccttccccacatctgt 1954
QY 2148 gatgcttgagtgacagagacaaactcccttggtgttccatactggtccttccataacatg 2207

Db 1955 gatgcttgagtgacagagacaaactcccttggtgttccatactggtccttccataacatg 2014
QY 2208 cctctcttccatcagtgcttcttgctgctgagctactgtggtgaagccttgccagaacta 2267
Db 2015 cctctcttccatcagtgcttcttgctgctgagctactgtggtgaagccttgccagaacta 2074
QY 2268 caacgagccaaatggtacacttccagctgtcttcaactcgtgtcctgatacagctt 2327
Db 2075 caacgagccaaatggtacacttccagctgtcttcaactcgtgtcctgatacagctt 2134
QY 2328 ctccacccgggccaagcgttccagcagcaggaagtaaccttgcctgcggccaacatgatactg 2387
Db 2135 ctccacccgggccaagcgttccagcagcaggaagtaaccttgcctgcggccaacatgatactg 2194
QY 2388 gctgagcagcctgagcagcagcctcgtgtggtatcttccgctaagtgtacgtatcct 2447
Db 2195 gctgagcagcctgagcagcagcctcgtgtggtatcttccgctaagtgtacgtatcct 2254
QY 2448 ctgcccggccagaccttccacacagagcacttccagccttccatccagagactacagag 2507
Db 2255 ctgcccggccagaccttccacacagagcacttccagccttccatccagagactacagag 2314
QY 2508 gcgctgcgctccacactga 2526
Db 2315 gcgctgcgctccacactga 2333

RESULT 6
AAD19503
ID AAD19503 standard; cDNA; 1101 BP.
XX
AC AAD19503;
XX
DT 18-DEC-2001 (first entry)
DE
DE Human novel G-protein coupled receptor (NGPCR) cDNA #3.
XX
XX Human: G-protein coupled receptor; GPCR; gene therapy; drug screening;
KW antisense-therapy; signal transduction; behavioural disorder; obesity;
KW heartbeats rate; inflammation; immune disorder; diabetes; cancer;
KW coronary disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1101
FT /tag= A
FT /product= "Human NGPCR protein"
FT unsure 520..522
FT /**tag= b
FT /note= "Encodes Ala"
XX
PD WO200172842-A2.
XX 04-OCT-2001.
XX 28-MAR-2001; 2001WO-US09996.
XX 28-MAR-2000; 2000US-192978P.
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A;
XX Zambrowicz B, Sands AT;
XX
XX WPI: 2001-616474/71.
XX P-PSDB; AAE11971.
XX
XX Novel isolated polynucleotides encoding human G protein coupled
XX receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
XX expressed NGPCRs for diagnosis of disease, and as probes or primers -

CC mutant NRCRs or inappropriately-expressed NRCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.

xx Sequence 705 BP; 130 A; 231 C; 170 G; 173 T; 1 other;

Query Match	27.98;	Score 705;	DB 22;	Length 705;
-------------	--------	------------	--------	-------------

Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1822	atgcttgagcctccctcgagagaagtaatggcaacgttaagagcttctttggggaaccaca	1881
Db	1	atgcttgagcctccctcgagagaagtaatggcaacgttctttggggaaccaca	60
QY	1882	agagctcgctgctgtgtacagccagagccctcttgcgcttggtttcacatcttcgtcc	1941
Db	61	agagctcgctgctgtgtacagccagagccctcttgcgcttggtttcacatcttcgtcc	120
QY	1942	tgcctgagcagttgcgtcatctcacatcatcatcttcaagtttccacaagaagtact	2000
Db	121	tgcctgagcagttgcgtcatctcacatcatcatcttcaagtttccacaagaagtact	180
QY	2002	acattctacacagcgtctggtgtgtccaaaacacacaggtgtgcgctcttggtagatcacgtcca	2061
Db	181	acattctacacagcgtctggtgtgtccaaaacacacaggtgtgcgctcttggtagatcacgtcca	240
QY	2062	ggcgcccaagctgcgtctatactgtcttaacttgcggtggtgtgtgaaacccaatgcgtctgata	2121
Db	241	ggcgcccaagctgcgtctatactgtcttaacttgcggtggtgtgtgaaacccaatgcgtctgata	300
QY	2122	gaatacacagcgtctcccccacatctgcgtgtgagtgtgacagagacacaactccctgtgac	2181
Db	301	gaatacacagcgtctcccccacatctgcgtgtgagtgtgacagagagacaactccctgtgac	360
QY	2182	ttaactatgagccttcctctctacaaatgagcctcctctccatcagtgcctttgcgtcagaactac	2241
Db	361	ttaactatgagccttcctctctacaaatgagcctcctctccatcagtgcctttgcgtcagaactac	420
QY	2242	cttggttaagaagactgtgcagagagaactacaaacagagagccaatgtgtataccttcaagcctgtgc	2301
Db	421	cttggttaagaagactgtgcagagagaactacaaacagagagccaatgtgtataccttcaagcctgtgc	480
QY	2302	ttaactctgtctctctgtgatacgcgctctctcttccacacacagcagagtgcttacaagcagaagtac	2361
Db	481	ttaactctgtctctctgtgatacgcgctctctcttccacacacagcagagtgcttacaagcagaagtac	540
QY	2362	ctgagcttcgagccaactatgatgtgtgtgcgtgcagagcgttaagcagaagcgtcttcggtgtgat	2421
Db	541	ctgagcttcgagccaactatgatgtgtgtgcgtgcagagcgttaagcagaagcgtcttcggtgtgat	600
QY	2422	tcttcgtcctaagtgctcagtaacatccctctgcgcgcacagaaacttaacagacagaagacttc	2481
Db	601	tcttcgtcctaagtgctcagtaacatccctctgcgcgcacagaaacttaacagacagaagacttc	660
QY	2482	cagagctccatctcaagatactacaagagagcgtctgcgtctccactgtga	2526
Db	661	cagagctccatctcaagatactacaagagagcgtctgcgtctccactgtga	705

RESULT	8
AAD17515	
ID	AAD17515 standard; DNA; 8191 BP.

XX		AAD17515;
AC		
XX		
DT	10-DEC-2001	(first entry)
XX		
DE		Human taste receptor, hTIR1 full-length genomic DNA.
XX		
KW		Human; taste-cell-specific G protein-coupled receptor; TIR1; drug;
KM		genetic modulation; pharmaceutical; taste sensation; food industry
RN		chemosensory transduction; ds.
XX		
OS		Homo sapiens.
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..7771
FT		/tag= a
FT		/product= "Human taste receptor, hTIR1 protein"
FT		1..846
FT	Intron	
FT		/tag= b
FT		/number= 1
FT	exon	/cons_splice= (5'site:NO, 3'site:NO)
FT		847..1037
FT		/tag= c
FT		/number= 1
FT	Intron	1038..1593
FT		/tag= d
FT		/number= 2
FT	exon	1594..1900
FT		/tag= e
FT		/number= 2
FT	Intron	1901..2819
FT		/tag= f
FT		/number= 3
FT	exon	2820..3581
FT		/tag= g
FT		/number= 3
FT	Intron	3582..4603
FT		/tag= h
FT		/number= 4
FT	exon	4604..4816
FT		/tag= i
FT		/number= 4
FT	Intron	4817..5138
FT		/tag= j
FT		/number= 5
FT	exon	5139..5259
FT		/tag= k
FT		/number= 5
FT	Intron	5260..6841
FT		/tag= l
FT		/number= 6
FT	exon	6842..7771
FT		/tag= m
FT		/number= 6
FT	Intron	7772..8191
FT		/tag= n
FT		/number= 7
FT		/cons_splice= (5'site:NO, 3'site:NO)
XX		
PN	WO20016563-A2.	
XX		
PD	13-SEP-2001.	
XX		
PF	07-MAR-2001; 2001WO-US07265.	
XX		
PR	07-MAR-2000; 2000US-0187546.	
PR	07-APR-2000; 2000US-0195536.	
PR	06-JUN-2000; 2000US-0209840.	
PR	23-JUN-2000; 2000US-0214213.	
PR	17-AUG-2000; 2000US-0226448.	
PR	03-JAN-2001; 2001US-0259227.	
PA	(SENO-) SENOMIX INC.	

CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. Inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.

CC Sequence 1147 BP: 249 A; 322 C; 331 G; 245 T; 0 other;

Query Match 21.3%; Score 538; DB 22; Length 1147;
Best Local Similarity 99.8%; Pred. No. 7.6e-247;
Matches 588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1006 gcgtttgaagaagcctatagcccgagcaagaagagcgcccttagccttgccacaagggc 1065
DB 89 gcgtttgaagaagcctatagcccgagcaagaagagcgcccttagccttgccacaagggc 148
QY 1066 tcctgtgacagcaagcaatcagctctgcagagaatgcagaagcttcatacgcacacagtg 1125
DB 149 tcctgtgacagcaagcaatcagctctgcagagaatgcagaagcttcatacgcacacagtg 208
QY 1126 cccaagctcaaaagccttcacatgattctgcctacaagcataccgggtgtgtatcg 1185
DB 209 cccaagctcaaaagccttcacatgattctgcctacaagcataccgggtgtgtatcg 268
QY 1186 gtggcccatggcctccaccagcctctggcgtgtgcctcttgagattgttccaggggcgca 1245
DB 269 gtggcccatggcctccaccagcctctggcgtgtgcctcttgagattgttccaggggcgca 328
QY 1246 gtctacccttgacgcttttgagacagatccacaaggtgcatttcctctacacaagagac 1305
DB 329 gtctacccttgacgcttttgagacagatccacaaggtgcatttcctctacacaagagac 388
QY 1306 actgtgacgtttaaatagacacagagatccctcagtagcctataacataattgtcctggac 1365
DB 389 actgtgacgtttaaatagacacagagatccctcagtagcctataacataattgtcctggac 448
QY 1366 tggagatgaccacaagtgagccttcaagctcctgcttcctcaatagttcccaagtttag 1425
DB 449 tggagatgaccacaagtgagccttcaagctcctgcttcctcaatagttcccaagtttag 508
QY 1426 ctaaacataaatagacccaataatccagtgacagcgaagaagcaaccaagtgctcctaagct 1485
DB 509 ctaaacataaatagacccaataatccagtgacagcgaagaagcaaccaagtgctcctaagct 568
QY 1486 gtgtgttcacagcagctgtcttgaaggcaccagcagtggttacgggttccatcaactgc 1545
DB 569 gtgtgttcacagcagctgtcttgaaggcaccagcagtggttacgggttccatcaactgc 628
QY 1546 tgccttagtggtgacccctggtggggtgggaccttcctccacaagagtg 1594
DB 629 tgccttagtggtgacccctggtggggtgggaccttcctccacaagagtg 677

RESULT 10

AAK68439 standard; DNA; 554 BP.

AAK68439;

06-NOV-2001 (first entry)

Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23251.
Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
PF 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.

XX Human immune/haematopoietic antigen encoding cDNA seq ID NO:2208.
DE
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
XX 09-AUG-2001.
PD 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0160628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.

PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234224.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234984.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-024617.
 PR 01-NOV-2000; 2000US-0246174.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPL: 2001-483426/52.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 24466; 3071bp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 1090 BP; 257 A; 283 C; 284 G; 266 T; 0 other;

Query Match 10.3%; Score 260; DB 22; Length 1090;
 Best Local Similarity 99.7%; Pred. No. 5.9e-114;
 Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 acaggtctgtagcttcaatagatgatagtctacacaccttccaaagtctagtgcttggg 247
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 753 acaggtctgtagcttcaatagatgatagtctacacaccttccaaagtctagtgcttggg 812
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 248 ttgaggagataaacaactccacggccctgctgcaccaatacacccttgggtaccagctgt 307
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 813 ttgaggagataaacaactccacggccctgctgcaccaatacacccttgggtaccagctgt 872
 QY 308 atgatgtgttctgacttgcacatgtatgtacacagctgagatgctctccctgcacg 367
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 873 atgatgtgttctgacttgcacatgtatgtacacagctgagatgctctccctgcacg 932
 QY 368 ggcaaccacacatagagcttcaaggagagaccttccactatccctcctaagtgctggag 427
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 933 ggcaaccacacatagagcttcaaggagagaccttccactatccctcctaagtgctggag 992
 QY 428 tgattggcctgacagacacacacgctgctgcacacacagcgcgcctgctgagccttcc 487
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 993 tgattggcctgacagacacacacgctgctgcacacacagcgcgcctgctgagccttcc 1052
 QY 488 tgggtgcccatg 498
 Db ||||||||||||
 Db 1053 tgggtgcccatg 1063

RESULT 13
 AAK69655 standard; DNA: 1090 BP.
 ID AAK69655
 XX

AC AAK69655;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:24467.
XX
KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229387.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 21-SEP-2000; 2000US-0234273.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0244676.
PR 08-NOV-2000; 2000US-0244677.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.

PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure: SEQ ID NO 24467; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (1)
XX amino acid sequences given in AAK62170 to AAK91921. (1) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (1)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (1) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (1) by expressing inactive proteins or to
XX supplement the patient's own production of (1). Additionally, (1)
XX polynucleotides may be used to produce the secreted (1), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (1) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK4703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK62169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 1090 BP; 258 A; 283 C; 283 G; 266 T; 0 other:
SO
Query Match 10.3%; Score 260; DB 22; Length 1090;
Best Local Similarity 99.7%; Pred. No. 5.9e-114;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 188 acaggtcttgtagctcaatgagcagtgctaccacctcttcagagctatgagcttgagg 247
DB 753 acaggtcttgtagctcaatgagcagtgctaccacctcttcagagctatgagcttgagg 812
OY 248 ttgagagataaacaacacacacgagccctgctgcccacaacacacacacacacacgt 307
DB 813 ttgagagataaacaacacacacgagccctgctgcccacaacacacacacacacacgt 872
OY 308 atgagtgtgtgtgtgactctgcacatgtgtatgcacagctgagagtgctctccctgcag 367
DB 873 atgagtgtgtgtgtgactctgcacatgtgtatgcacagctgagagtgctctccctgcag 932
OY 368 ggcac 427
DB 933 ggcac 992
OY 428 ttattgggctctacagac 487
DB 993 tgattgggctctacagac 1052
OY 488 tgggtgcccatg 498
DB 1053 tgggtgcccatg 1063
RESURF 14
ID ABA08327 standard; cDNA; 410 BP.

XX
AC ABA08327;
XX 11-JAN-2002 (first entry)
DT
XX
DE Human sensory GPCR B3 homologue-encoding cDNA, SEQ ID NO:103.
XX
XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; antiinflammatory;
XX antasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
XX cyostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
XX antifungal; vulnerary; antilucer; ss.
XX
XX Homo sapiens.
OS
XX WO200157188-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US03800.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YF, Liu C, Dermanac RT;
XX WPI, 2001-457740/49.
XX P-PSDB; ABB11083.
XX
XX Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer -
XX
XX Claim 1; Page 377; 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a
XX nucleotide of the invention, methods of producing the novel polypeptides,
XX antibodies against the polypeptides, methods of detecting the nucleotides
XX or polypeptides in a sample, and methods of identifying compounds which
XX bind to polypeptides of the invention. Although novel, many of the
XX polypeptides of the invention have homology to known proteins, thereby
XX giving an insight into their probable biological activities, and hence
XX potential therapeutic applications. The polypeptides of the invention may
XX have various activities, including cytokine, cell proliferation or cell
XX differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin- or inhibin-related activities;
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be
XX involved in oncogenesis, cancer cell proliferation or metastasis.
XX Depending on their biological activities, polypeptides and nucleotides of
XX the invention are useful for preventing, treating or ameliorating medical
XX conditions, e.g., by protein or gene therapy. Such conditions include
XX cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
XX proliferative retinopathy, atherosclerosis, coronary heart disease,
XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
XX vascular growth. Polypeptides involved with tissue regeneration and
XX repair (or nucleic acids encoding them) may be used to promote wound
XX healing (e.g., of burns, incisions and ulcers), while those with
XX immunomodulatory activities may be used in the treatment of viral,

CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 410 BP; 74 A; 135 C; 99 G; 102 T; 0 other;

Query Match 7.6%; Score 191; DB 22; Length 410;
Best Local Similarity 99.6%; Pred. No. 5.8e-81;
Matches 241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 247 gtgagagagtaaacactccagcgccctgtgcccacatcacctgggtaccagctg 306
DB 147 gtgagagagtaaacactccagcgccctgtgcccacatcacctgggtaccagctg 206
QY 307 tatgatgtgtgtgtgactctgccaatgtgtatgccacgctgagagtgctccctccca 366
DB 207 tatgatgtgtgtgtgactctgccaatgtgtatgccacgctgagagtgctccctccca 266
QY 367 gggcaaacaccacatagagctccaagagacctctccactatccctacagtgctggca 426
DB 267 gggcaaacaccacatagagctccaagagacctctccactatccctacagtgctggca 326
QY 427 gtgatgtggctgacagcaccacacgctgctgccacacacgcccctgctgagcccttc 486
DB 327 gtgatgtggctgacagcaccacacgctgctgccacacacgcccctgctgagcccttc 386
QY 487 ct 488
DB 387 ct 388

RESULT 15
AAZ50743
ID AAZ50743 standard; cDNA; 2771 BP.
XX
AC AAZ50743;
XX
DT 31-MAY-2000 (first entry)
XX
DE Rat sensory transduction G-protein coupled receptor-B3 cDNA.
XX
KW Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
KW sensory cell; taste receptor cell; screen; taste modulator;
KW pharmaceutical; food industry; taste topographic map; tongue; ss.
XX
OS Rattus sp.
XX
FH key Location/Qualifiers
FT CDS 57..2579
FT /tag- a
FT /product- "GPCR-B3"
XX
PN MO200006592-A1.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-US17099.
XX
PR 28-JUL-1998; 98US-0094465.
XX
PA (REGC) UNIV CALIFORNIA.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
XX WPI; 2000-205451/18.
DR

DR P-PSDB; AAY45021.
XX
CC New isolated sensory transduction G-protein coupled receptor, useful
PT for developing products for use in studying and modulating the taste
PT transduction pathway -
XX
PS Claim 5; Pages 76-77; 83pp; English.
XX
CC The present sequence is a cDNA encoding a taste cell specific G-protein
CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
CC This sequence was isolated from the 12AP rat circumvallate cDNA library.
CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
CC lower expression in circumvallate taste receptor cells of the tongue.
CC The present sequence is used to screen compounds that modulate sensory
CC signalling in taste cells, especially taste modulators useful in
CC pharmaceutical and food industries to customise taste. The sequence
CC can also be used as probe for identifying taste cells and
CC subsets of taste receptor cells such as foliate, fungiform and
CC circumvallate. Such probes are also useful to generate taste
CC topographic maps that elucidate the relationship between the taste
CC cells of the tongue and sensory neurons leading to taste centres
CC in the brain.
XX
SQ Sequence 2771 BP; 564 A; 797 C; 754 G; 656 T; 0 other;

Query Match 1.3%; Score 32; DB 21; Length 2771;
Best Local Similarity 100.0%; Pred. No. 5.6e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2119 agggataaccagcgcttcccccatctgtgat 2150
DB 2172 agggataaccagcgcttcccccatctgtgat 2203

Search completed: June 11, 2002, 09:19:48
Job time: 5800 sec